

AD-A239 817



2

NCS TIB 90-6



NATIONAL COMMUNICATIONS SYSTEM

TECHNICAL INFORMATION BULLETIN 90-6

STANDARDIZATION OF END-TO-END PERFORMANCE FOR FULL MOTION VIDEO TELECONFERENCING

MAY, 1990



OFFICE OF THE MANAGER
NATIONAL COMMUNICATIONS SYSTEM

WASHINGTON, D.C. 20305

91-07835



DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

91 8 14 014

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page instructions for filling in each block of the form follow. It is important to stay *within the lines* to meet optical scanning requirements.

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered
State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	PR - Project
G - Grant	TA - Task
PE - Program Element	WU - Work Unit Accession No.

Block 6. Author(s) Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes Enter information not included elsewhere such as Prepared in cooperation with..., Trans. of..., To be published in... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement
Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.

NASA - Leave blank.

NTIS - Leave blank.

Block 13. Abstract. Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

Block 14. Subject Terms Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS only).

Blocks 17 - 19. Security Classifications Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

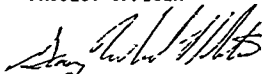
Block 20. Limitation of Abstract This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

NCS TECHNICAL INFORMATION BULLETIN 90-6

STANDARDIZATION OF END-TO-END PERFORMANCE
FOR FULL MOTION VIDEO TELECONFERENCING

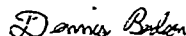
MAY 1990

PROJECT OFFICER



GARY REKSTAD
Electronics Engineer
Office of NCS Technology
and Standards

APPROVED FOR PUBLICATION:



DENNIS BODSON
Assistant Manager
Office of NCS Technology
and Standards

FOREWORD

Among the responsibilities assigned to the Office of the Manager, National Communications System, is the management of the Federal Telecommunication Standards Program. Under this program, the NCS, with the assistance of the Federal Telecommunication Standards Committee identifies, develops, and coordinates proposed Federal Standards which either contribute to the interoperability of functionally similar Federal telecommunication systems or to the achievement of a compatible and efficient interface between computer and telecommunication systems. In developing and coordinating these standards, a considerable amount of effort is expended in initiating and pursuing joint standards development efforts with appropriate technical committees of the Electronics Industries Association, the American National Standards Institute, the International Organization for Standardization, and the International Telegraph and Telephone Consultative Committee of the International Telecommunication Union. This Technical Information Bulletin presents an overview of an effort which is contributing to the development of compatible Federal, national, and international standards in the area of Video Teleconferencing. It has been prepared to inform interested Federal activities of the progress of these efforts. Any comments, inputs or statements of requirements which could assist in the advancement of this work are welcome and should be addressed to:

Office of the Manager
National Communications System
ATTN: NCS-TS
Washington, DC 20305-2010

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



STANDARDIZATION OF END-TO-END
PERFORMANCE FOR FULL MOTION
VIDEO TELECONFERENCING

May 21, 1990

SUBMITTED TO:
NATIONAL COMMUNICATIONS AGENCY
Office of Technology and Standards
WASHINGTON, D.C. 20305

Contracting Agency:
DEFENSE COMMUNICATIONS AGENCY
Contract Number - DCA100-87-C-0078
Task Order Number - 87-010

DELTA INFORMATION SYSTEMS, INC
Horsham Business Center, Bldg. 3
300 Welsh Road
Horsham, PA 19044

TEL: (215) 657-5270

FAX: (215) 657-5273

TABLE OF CONTENTS

SECTION 1 - INTRODUCTION AND SUMMARY	1 - 1
SECTION 2 - AVAILABLE BACKGROUND MATERIAL	2 - 1
2.1 Previous NCS Program	2 - 1
2.2 Testing of Compressed Video Signal Transmission	2 - 1
SECTION 3 - STILL PICTURE MEASUREMENTS	3 - 1
3.1 General Approach	3 - 1
3.2 Performance Parameters	3 - 3
3.3 Instrumentation	3 - 6
3.4 Proposed Performance Standards	3 - 23
SECTION 4 - MOVING PICTURE MEASUREMENTS	4 - 1
4.1 Background	4 - 1
4.2 Scene Cut Patterns	4 - 3
4.3 Rotating Wheel Patterns	4 - 7
4.4 Test Tape Preparation	4 - 12
4.5 Measurement Implementation	4 - 14
SECTION 5 - MOTION TEST METHODOLOGY ANALYSIS	5 - 1
5.1 Data Collection	5 - 1
5.2 Numerical Evaluation	5 - 2
5.3 Results	5 - 6
SECTION 6 - CONCLUSIONS AND RECOMMENDATIONS	6 - 1
6.1 Summary of Results	6 - 1
6.2 Recommended Further Still Picture Analysis	6 - 2
6.3 Recommended Further Motion Analysis	6 - 3

LIST OF ILLUSTRATIONS

Figure 3 - 1	Combination Test Signal - Modified	3 - 11
Figure 3 - 2	Combination Test Signal - Modified	3 - 12
Figure 3 - 3	Combination Test Signal - Modified	3 - 13
Figure 3 - 4	Composite Test Signal - Modified	3 - 14
Figure 3 - 5	Shallow Ramp - Low Level	3 - 15
Figure 3 - 6	Shallow Ramp - Medium Level	3 - 16
Figure 3 - 7	Shallow Ramp - High Level	3 - 17
Figure 3 - 8	Multi-step Flat Bottom	3 - 18
Figure 3 - 9	Multi-step Flat Top	3 - 19
Figure 3 - 10	Color Bar Chart Test Signal	3 - 21
Figure 3 - 11	Field Bar Test Signal	3 - 22
Figure 4 - 1	Circle Pattern	4 - 5
Figure 4 - 2	30 Degree Wheel Pattern	4 - 9
Figure 4 - 3	Test Setup for Motion Performance Evaluation	4 - 16
Figure 5 - 1	Test Data Recording Form	5 - 4
Figure 5 - 2	Waveform & Computed RMS Value	5 - 5
Figure 5 - 3	Spoke Size Comparison - 1544 KBPS	5 - 11
Figure 5 - 4	Spoke Size Comparison - 384 KBPS	5 - 12
Figure 5 - 5	Spoke Size Comparison - 56 KBPS	5 - 13
Figure 5 - 6	Bit Rate Comparison - 18° Spokes	5 - 14
Figure 5 - 7	Bit Rate Comparison - 10° Spokes - Codec L	5 - 15
Figure 5 - 8	Bit Rate Comparison - 10° Spokes - Codec C	5 - 16
Figure 5 - 9	Codec Comparison - 18° Spokes	5 - 17
Figure 5 - 10	Codec Comparison - 10° Spokes	5 - 18
Figure 5 - 11	Codec Comparison - 30° Spokes	5 - 19
Figure 5 - 12	Circle Size Comparison - 384 KBPS	5 - 20
Figure 5 - 13	Circle Size Comparison - 56 KBPS	5 - 21
Figure 5 - 14	Bit Rate Comparison - Small Circles	5 - 22
Figure 5 - 15	Bit Rate Comparison - Medium Circles	5 - 23
Figure 5 - 16	Codec Comparison - Medium Circles	5 - 24

LIST OF TABLES

Table 3 - 1	Codec Test Signals	3 - 20
Table 3 - 2	Waveform Parameter Standards and Limits	3 - 24
Table 3 - 3	Performance Parameter Standards and Limits	3 - 26
Table 4 - 1	Scene Cut Patterns	4 - 6
Table 4 - 2	Rotating Wheel Parameters	4 - 10
Table 4 - 3	Motion Arrangements	4 - 13
Table 5 - 1	Temporal Frequencies	5 - 7

SECTION 1 INTRODUCTION AND SUMMARY

This document summarizes work performed by Delta Information Systems, Inc. (DIS) for the National Communications System (NCS), Office of Technology and Standards. This office is responsible for the management of the Federal Telecommunications Standards Program, which develops telecommunications standards, whose use is mandatory for all Federal departments and agencies. This study was performed under task order number 87-010 of contract number DCA100-87-C-0078.

This report covers the standardization of end-to-end performance for full motion video teleconferencing. It investigates the feasibility of such a standardization on an end-to-end (camera-to-monitor) basis to ensure delivery of a specified quality of service by developing a candidate set of objective measures of TV picture quality for Government TV applications. This is a distinct difference from currently used methods which depend entirely on subjective evaluations.

Objective still picture measurements are standard in the broadcast industry and with other analog video systems. These methods can be adapted to digital video teleconferencing. However, there are no available techniques for the objective analysis of moving pictures which is a vital element of the performance evaluation of any digital video teleconferencing system. Development of such techniques is a major objective of the effort covered by this report.

Section 2 briefly describes previous programs which have provided background material for the current effort. The proposed performance standards based on still picture measurements are contained in Section 3. The methods and test patterns which were developed for experimental measurements on moving pictures are described in Section 4. They lead to the preparation of a special motion test tape and development of a method for the performance of measurements on this tape after processing through a digital video codec. Section 5 discusses the analysis of these results and develops groups of preliminary performance curves which show the validity of the developed methodology. Section 6 summarizes the results of the programs and makes recommendations for related future efforts.

SECTION 2 - AVAILABLE BACKGROUND MATERIAL

2.1 Previous NCS Program

A first step towards the objectives of this program is documented in the report on the Development of an Objective Standardized Test for Video Teleconferencing, dated February, 1987. That report covered an early approach to the very ambitious goal of establishing objective test methods for digital video codecs. The study depended mainly on standard video tests signals recorded on the video tape used for subjective tests which was processed through four different codecs. The same tape also contained an attempt to simulate motion with a fast pixel change produced by switching between two well defined pictures.

The results of this program fell short of its stated goal of achieving correlation of subjective and objective test results. First of all, taping the test signals produced too much inherent distortion to provide a dependable input signal. Furthermore, many of the conventional analog parameters proved to be of little importance to the subjective observers because their evaluations were mainly based on the motion performance determined by the digital algorithm of the codec. Nevertheless, the program provided good understanding of the many applicable criteria and inherent problems and pointed the way for future studies. Avoidance of the shortcomings which manifested themselves furnished valuable guidance for the present effort.

2.2: Testing of Compressed Video Signal Transmission

This program was performed for the Defense Communication Agency (DCA) as a subcontract to TRW. It is documented in three reports dated in September, October and December 1985. The purpose of the program was the establishment of objective standards and measurement techniques for determining compressed video signal network performance for networks operating at a data rate of 1.544 Mbps. The reports specifically address the codecs of one manufacturer, Compression Labs, Inc. (CLI). This was done since no standard data rate or compression algorithm had been officially established, but many U.S. Government and commercial installations had selected the CLI VTS-1.5E codec; therefore, a CLI codec was felt to be representative.

It is obvious that the objectives of this past program, though much more limited, are fairly similar to the present effort for NCS. The most important results are the analog parameters that should be measured in a digital video system, their standard values and permissible performance limits. These values were backed up by a series of equipment tests. With some modifications, the results will be applicable to the full range of codecs and data rates to be covered in the present program.

SECTION 3 - STILL PICTURE MEASUREMENTS

3.1 General Approach

The end-to-end performance of any video system requires the measurement of a number of still picture parameters. This number will vary depending on the picture quality requirements of the specific system. For instance, only minimal measurements are needed in a home video system, while broadcasting and cable distribution systems must satisfy very high standards. Evaluation of motion performance of analog video systems generally is not necessary since the old problem of "sticky" camera tubes has been virtually eliminated with modern video pickup devices.

The methods for testing digital video systems and developing performance standards are presently going through similar steps as broadcast TV in the 1950's. Evaluation of picture quality was then achieved subjectively by many panels of experts. Simultaneously, an ever increasing list of performance parameters deemed likely to affect picture quality was compiled and the signal waveforms were standardized. Correlation of objective parameter measurements with subjective evaluation resulted in the selection of meaningful performance parameters and establishment of limit values for satisfactory picture quality.

The many parameters that describe the format and fidelity of a color video signal are contained in three well established official documents. Electronic Industries Association (EIA)

Standard RS-170A specifies signal waveforms, amplitudes and timing. EIA Standard RS-250B and Network Transmission Committee (NTC) Report No. 7 are similar in content and specify acceptable limits of signal degradation after passage through a transmission system. Since all these documents date back to the 1976/77 time period, some added parameters have become popular since then. Some limited information was obtained from portions of a draft version of RS-250C but this document is still under committee consideration, therefore the contents of the draft are not official. No significant changes are anticipated.

The American National Standards Institute (ANSI) has recently issued Standard T1-502 for the transmission of high quality analog video signals. This standard features several new performance parameters and much tighter limits for most previously used parameters, which are not realistic for low data rate digital video codecs. Committee T1Q1.5 is in the process of adapting T1-502 for the digital transport of video teleconferencing/video telephony signals but more work remains to be done. The NTC has drafted an updated report NTC-9 but suspended its effort before publication. CCIR Recommendation 567 is applicable to long international connections and therefore has rather lenient performance requirements. Consequently, RS-250B and NTC-7 still provide the most useful background material for the present program. Recently published cable system (CATV) performance specifications are largely based on the same documents.

The long lists of parameters and performance specifications in these two documents present a logical point of departure for the development of performance standards of a video teleconferencing system. The characteristics and applications of a digital system affect the importance and acceptable limit values of the conventional analog parameters. While some of them can probably be completely eliminated, the characteristics of a digital system may make it desirable to add some measurements to fully evaluate system performance.

3.2 Performance Parameters

The first step in the development of standards of video teleconference systems is the selection of the necessary parameters. This can best be accomplished by reviewing existing analog TV standards, with due consideration of inherent differences between analog and digital video transmissions and equipments. A major factor is that a video codec transmits the input signal not in its actual form but only enough digital information to allow regeneration of the analog signal at the receiver.

Any video signal that is to be displayed on a standard NTSC monitor must basically comply with RS-170A, though many amplitude and timing limits can be relaxed. However, codecs generally deviate considerably from the specified horizontal and vertical blanking widths. This is done to allow codecs to economize in the amount of transmitted data by slightly reducing the picture

size in both directions without eliminating useful information.

The codec re-constitutes the analog signal from the compressed digital format in which it is transmitted. Signal waveforms and timing are generated in the decoder and therefore not directly dependent on the same parameters at the point of entry into the encoder. They are also not affected by the encoding/decoding process and the transmission circuit. Since generation of analog TV signals is a routine matter, a decoder design compliant with RS-170A presents no problem.

A much closer scrutiny is required of the parameters contained in RS-250B and NTC Report No. 7. They affect picture quality in a known and readily visible fashion. It is obviously advisable to utilize the many years of experience of the broadcast industry, but the characteristics of a video codec dictate some variations. A few parameters which are mainly affected by certain factors typical of analog transmission become largely irrelevant in a digital transmission system. Non-linear transfer characteristics and dynamic gain distortions are often caused by limitations in FM detectors and low frequency response which do not affect digital transmission. Therefore, measurements of dynamic gain, long-time waveform distortion (bounce), and use of average picture levels (APL) other than 50% in differential gain and phase measurements become unnecessary. Transmission noise of all types is highly unlikely to affect the received picture because the decoder is tolerant to error rates of up to 10^{-6} before forward error correction. This allows noise

levels much higher than those acceptable in analog transmission. Error rates above 10^{-6} cause very noticeable and rapidly increasing degradations which soon result in complete system failure. The only noise possibly noticeable at the decoder output is a sum of quantizing noise and contributions from power supplies and other portions of the codec output circuit. This noise level is so low that it is difficult to obtain consistent and repeatable results, therefore a value could not yet be specified. Limited results obtained so far indicate that a value in the range of -60 dB is probably appropriate. The output level of the re-constituted signal (often called insertion gain), once properly set, is unlikely to vary. Field-time waveform distortion caused by low frequency response limitations is likely to be low and constant.

The practical experience gained by the codec performance measurements made as part of the programs described in Section 2 corroborates the above statements. Beyond that, it shows that at 1.544 Mbps most conventional performance parameters have values either fully or almost complying with analog specifications. When trying to correlate the measurement results with the subjective evaluations it became apparent that in most cases little or no correlation existed. The only significant parameter was frequency response which very obviously affected the received picture quality. Recent measurements made on three codecs have shown considerable differences in the range between 1 and 3 MHz, with 3 MHz being the highest value worth specifying. Exact

values over the whole range still remain to be determined. Different specifications for high and low bit rates are expected to be needed. This is due to the fact that frequency response is a necessarily severe limitation in video codec performance when compared with analog transmission.

A popular pattern for a quick overall check of video system performance is the color bar chart. It shows the approximate combined effect of various degradations in color performance parameters. Therefore, a parameter called "vector accuracy", giving the amplitude and phase deviations of six color vectors, becomes a very convenient and useful tool in video teleconferencing system performance evaluation.

A feature unique to digital video systems is the accuracy of sampling and quantization in the analog/digital conversion process. Experience in this area is limited and neither methods nor performance values for digital video have been suggested. It is important that this parameter and signals for its evaluation be defined in order to establish complete standards of video teleconferencing systems. Recent preliminary tests with experimental test signals have not yet achieved results which are consistent enough to form the basis for a proposed test procedure and specification.

3.3 Instrumentation

The measurements performed for the previous NCS program used the test signal portion of the 1" test tape prepared for

subjective codec evaluation. After processing through the four codecs under test, many parameters showed high values of distortion. Subsequent measurement on the original test tape showed that a considerable portion of these distortions was present at the input. At that time it could not be determined whether this was due to the input test signal or the taping process. At a later time test signals known to be perfect were recorded on tape which again produced distorted signals.

The results in the previous NCS program were obtained as the difference between codec input and output distortion values. This method cannot be considered accurate because these distortions are not necessarily arithmetically additive. Nevertheless, the results proved to be logical and usable. One fact that became obvious is that parameters such as differential gain and phase should be measured only at an average picture level (APL) of 50%. The conventional test signals for APL values of 10% and 90% have color content only on every fifth line. This feature, in conjunction with the color signal line rate subsampling used in all codecs, produces such a low output signal that it becomes too noisy to be properly measured. Fortunately, as stated in Paragraph 3.2, measurements at 50% APL are sufficient because transmission non-linearities do not affect the codec output. However, in summary, it must be stated that tape recorded test signals cannot be recommended.

During the DCA/TRW program, high quality modern test equipment was made available at the codec location. Therefore,

all measurements could be made directly and the results are dependable and accurate. The tests were implemented with the Tektronix 1910 NTSC Digital Signal Generator and the 1980 Automatic Video Measurement Set ("Answer"). The equipment was configured for the analysis of broadcast quality signals and printouts of the test results highlighted all deviations from the pre-set narrow limits. In spite of that, more than half of all measured parameters were found compliant. Obvious and necessary deviations were horizontal and vertical blanking widths, which are deliberately increased in the codec, and frequency response, with related parameters in both luminance and chrominance channels. These results showed that some test signals should be modified to be better adapted to the performance limitations of a codec.

These previous results clearly show that objective still picture measurements should be made only directly with co-located codec and test equipment. DIS has meanwhile obtained a 1910 Signal Generator and VM-700 Video Measurement Set which is the much improved successor to the 1980 "Answer". The VM-700 includes a display which can be used for viewing all waveforms and several programmed sets of performance data before print-out. Since the 1910 is an all-digital signal generator its output signals can be changed by inserting re-programmed PROM's. Such PROM's have been procured for the DIS unit to investigate codec testing by substituting seven different experimental signals.

Two normally available signals have been modified. The

three high frequency packets in the multiburst portion of the NTC-7 combination test signal have been changed to measure a more limited frequency response. The chrominance pulse of the NTC-7 composite test signal has been changed from a half-amplitude duration of $12.5 T$ (1562.5 nsec) to $20T$ (2500 nsec) to accommodate the limited bandwidth of codec chrominance circuits. The five other signals are new, and mainly for the purpose of developing a method to evaluate the quantizing performance of the analog/digital conversion process. A shallow ramp signal is useful to check and possibly measure quantizing errors. Therefore, three of the provided signals consist of ramps with an amplitude of 16 IRE units superimposed on pedestals of 18, 50 and 92 IRE units to investigate the quantization performance at low, medium and high picture levels. The other two signals consist of 12 multiple-amplitude steps, one each with constant minimum or maximum level. These signals may be useful to investigate effects such as overshoots of non-linear DPCM near both the white and black levels.

Figures 3-1 to 3-9 show the modified signals, including the new packet frequencies as plotted on a VM-700 Video Measurement Set. The signals of Figure 3-1 to 3-4 will be used in conventional fashion to measure the appropriate video parameters. The exact application and related numerical evaluation of the other signals still has to be determined experimentally. Recently gained experience has shown some usefulness of the shallow ramps in determining quantizing accuracy but more effort

is needed. Further changes of the modified test signals may prove to be desirable. It must be emphasized that these modifications do not interfere with the normal use of the 1910 Signal Generator which is readily restored by changing back to the original PROM's.

The performance parameters and required test signals are shown on Table 3-1. All except the last two signals are included in the combination and composite test signals already shown. The remaining two signals are standard and shown on Figures 3-10 and 3-11. No special test signal is required for measurement of waveform parameters because the synchronizing signal is always present.

All these test signals occupy a full field. This is inherently necessary for the field bar signal but all the other signals can be contained in a single line. It has become standard broadcast practice to insert test signals on one or several lines during the vertical blanking, creating vertical interval test signals (VITS) which allow continuous in-service monitoring. This cannot be done with a codec signal because the vertical interval is not transmitted. Should in-service monitoring become desirable in a future video teleconference system, test signals would have to be inserted on lines at the top and/or bottom of the picture. If it is possible to preempt those lines whenever needed, such "pseudo-VITS" could be established as part of a future digital video standard.

VM700 Video Measurement Set
Channel A Channel A 12-May-88 15:45:26

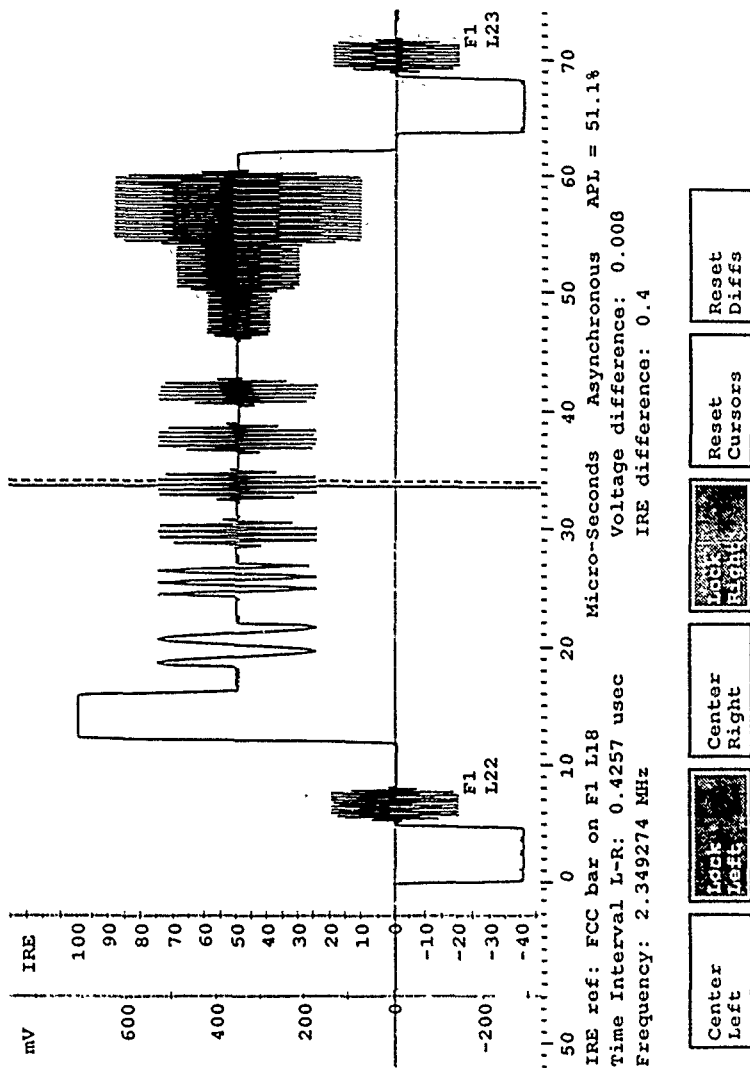


FIGURE 3-1: COMBINATION TEST SIGNAL - MODIFIED - SHOWING 2.35 MHZ BURST

VM700 Video Measurement Set

Channel A Channel A 12-May-88 15:39:39

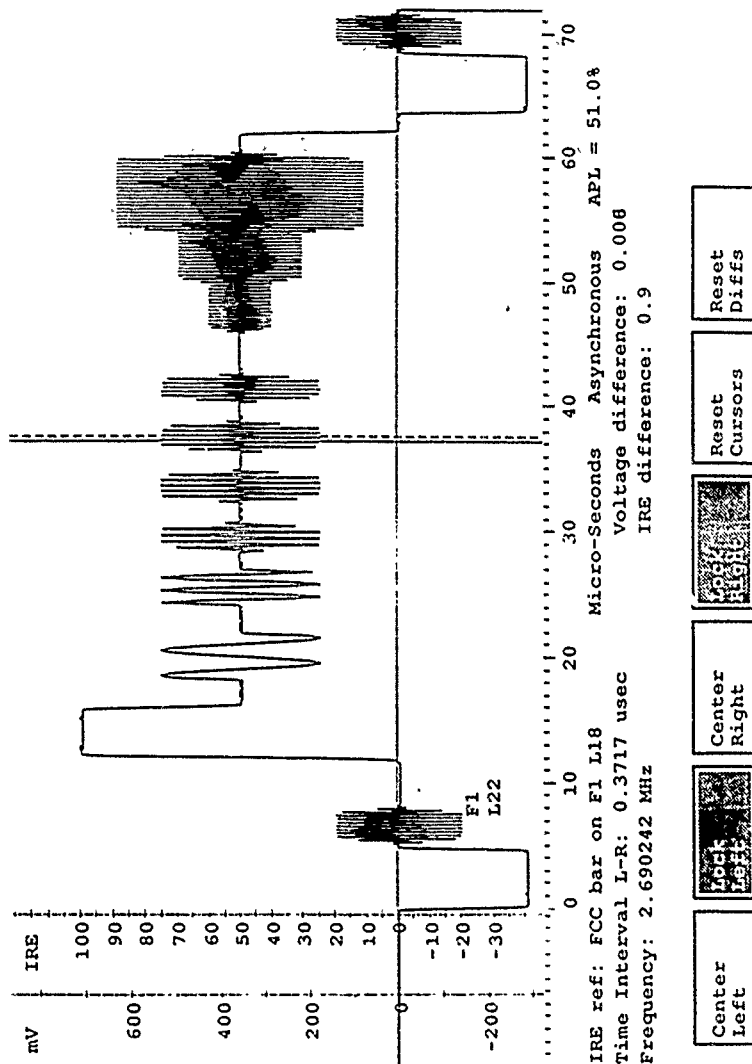


FIGURE 3-2: COMBINATION TEST SIGNAL - MODIFIED - SHOWING 2.7 MHZ BURST

VM700 Video Measurement Set

Channel A Channel A 12-May-88 15:35:49

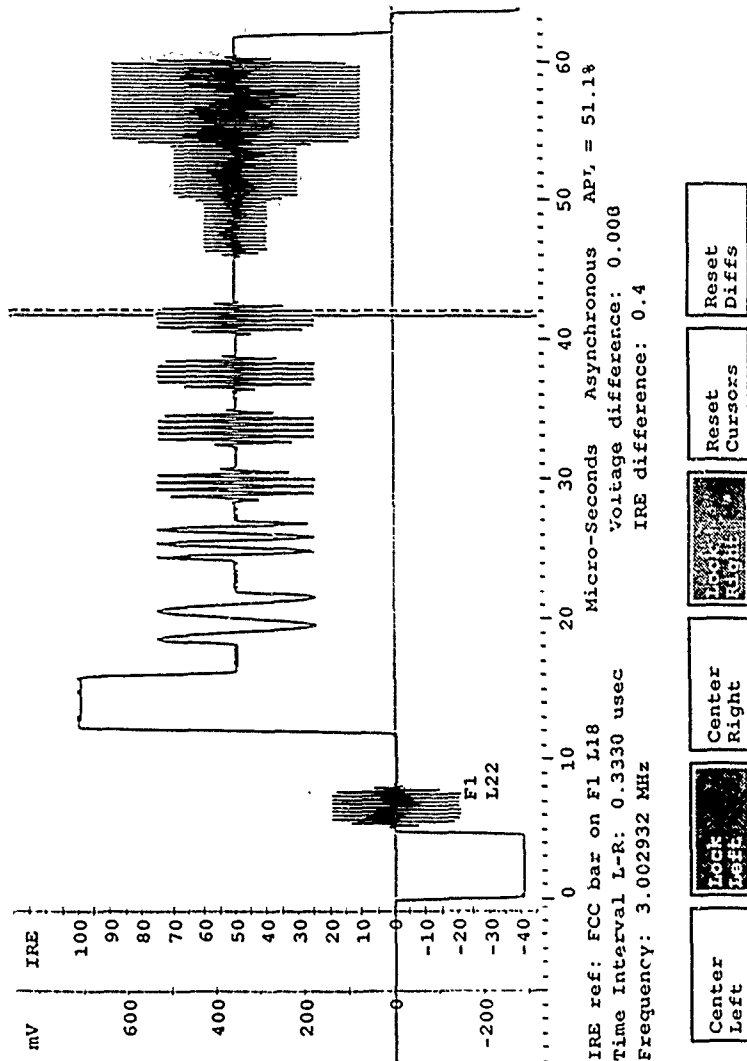


FIGURE 3-3: COMBINATION TEST SIGNAL MODIFIED - SHOWING 3.0 MHz BURST

VM700 Video Measurement Set

Channel A Channel A

12-May-88 15:13:04

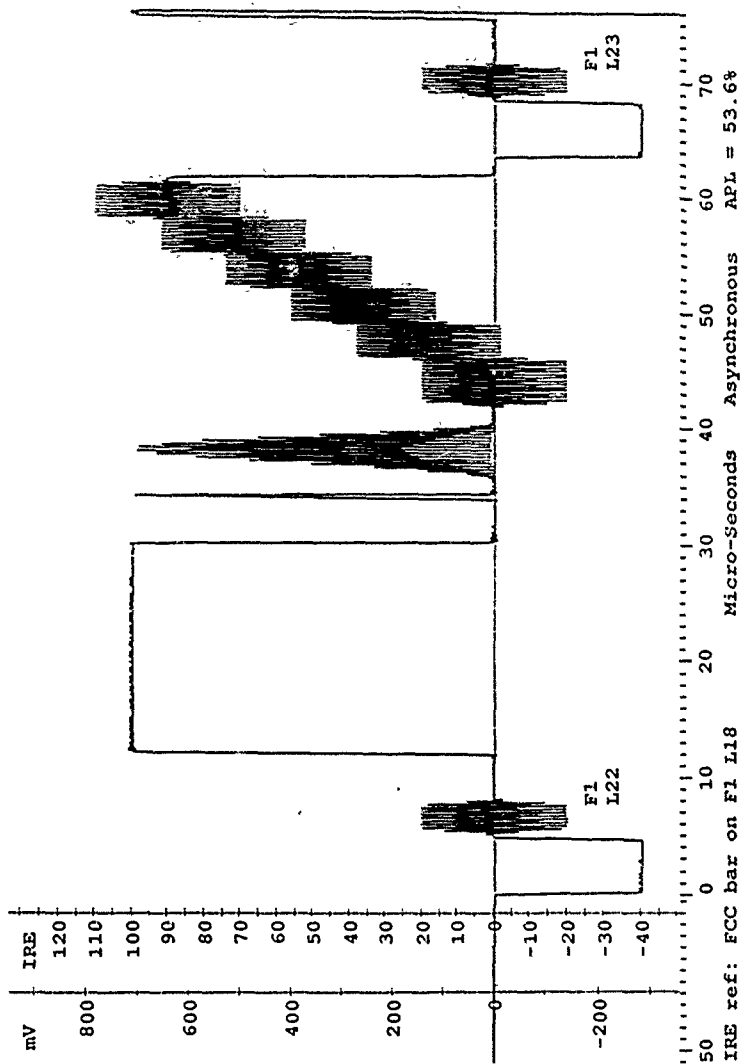


FIGURE 3-4 COMPOSITE VIDEO SIGNAL - ANALOG

VM700 Video Measurement Set

Channel A Channel A 12-May-88 15:16:09

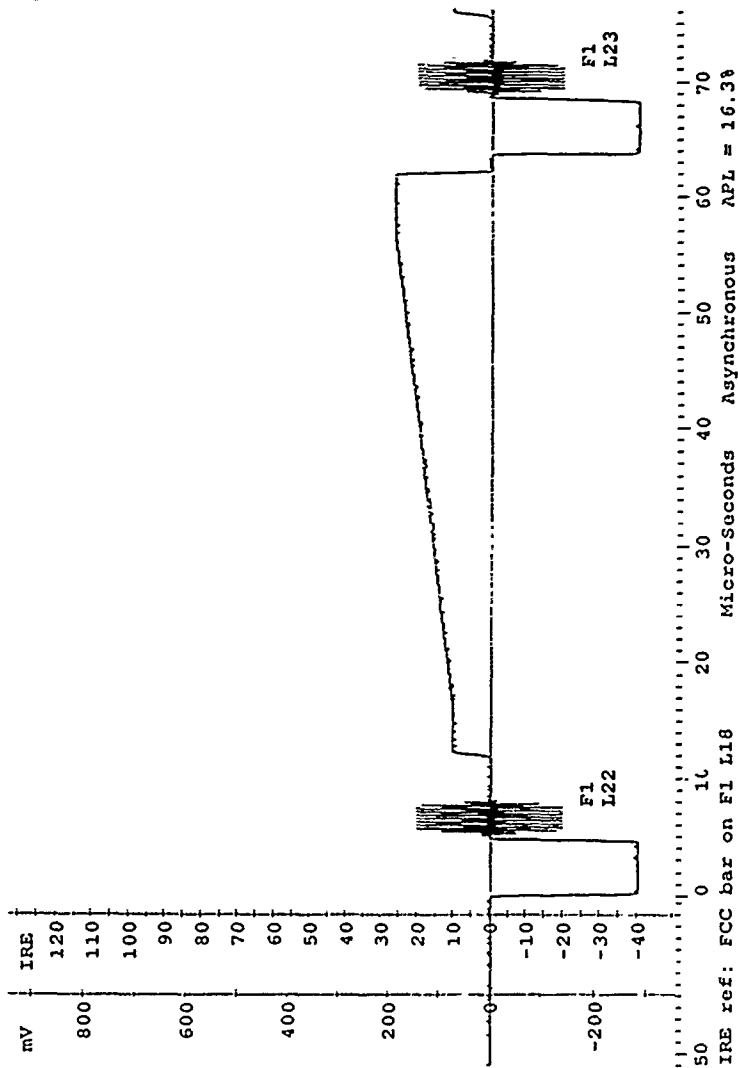


FIGURE 3-5 SHALLOW, CAMP - LOW LEVEL

VM700 Video Measurement Set
Channel A Channel A 12-May-88 15:17:09

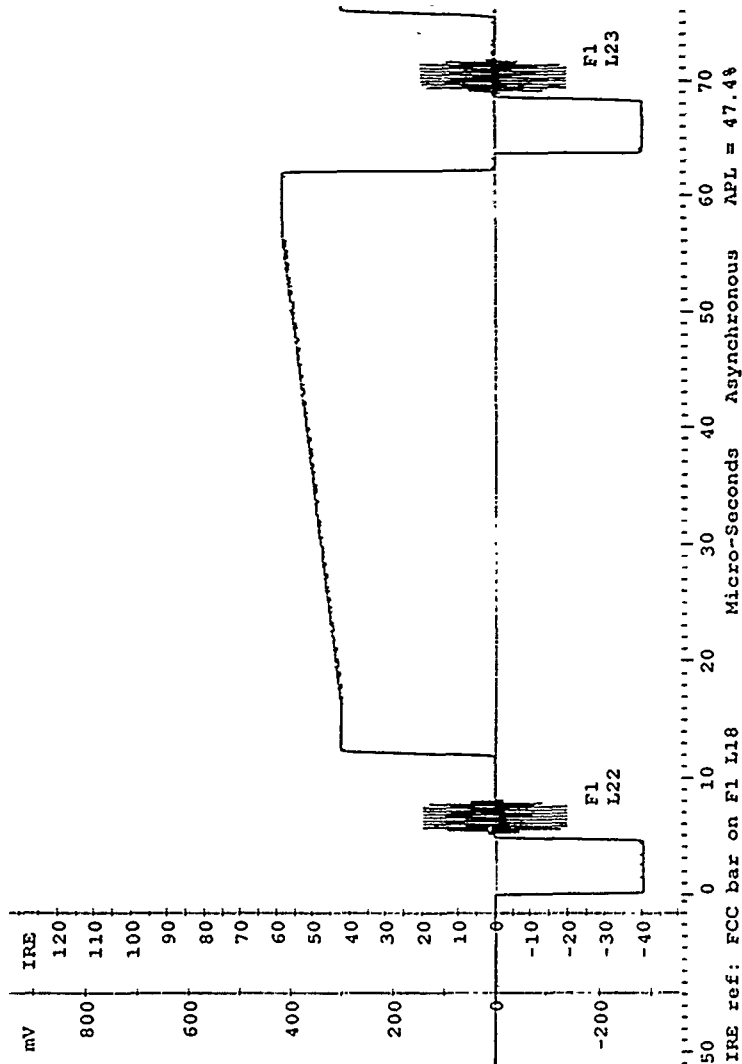
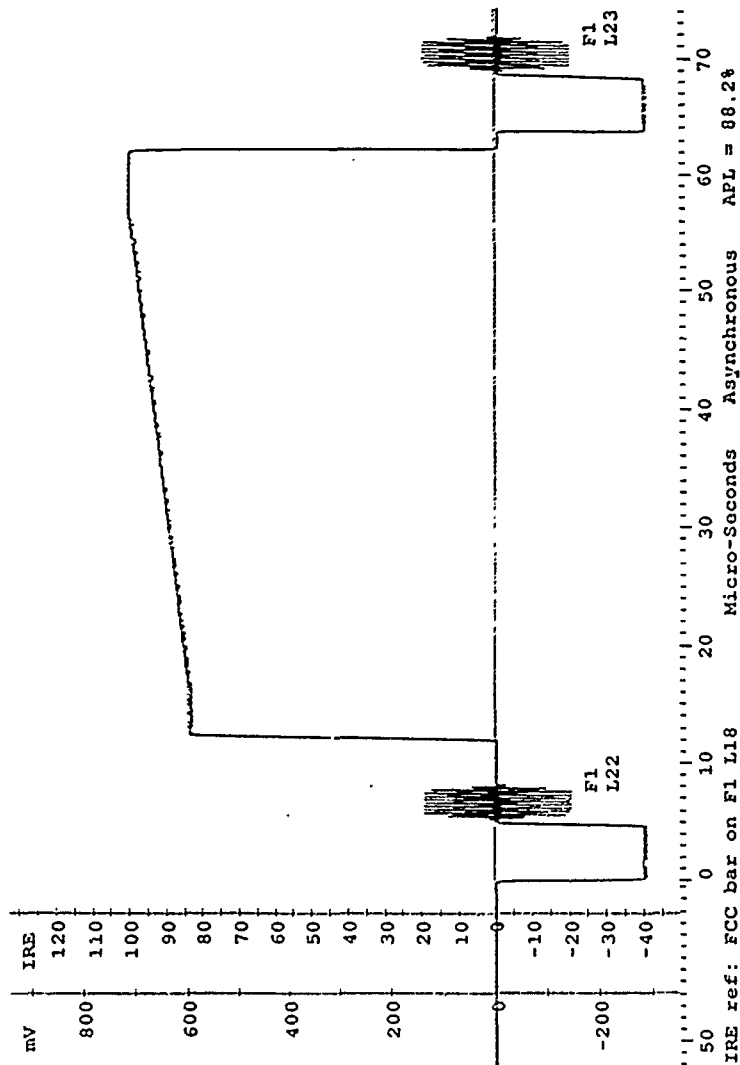


FIGURE 3 - 6 SHALLOW RAMP - MEDIUM LEVEL

VM700 Video Measurement Set
 Channel A Channel A 12-May-88 15:50:49



VM700 Video Measurement Set

Channel A Channel A

12-May-88 15:52:03

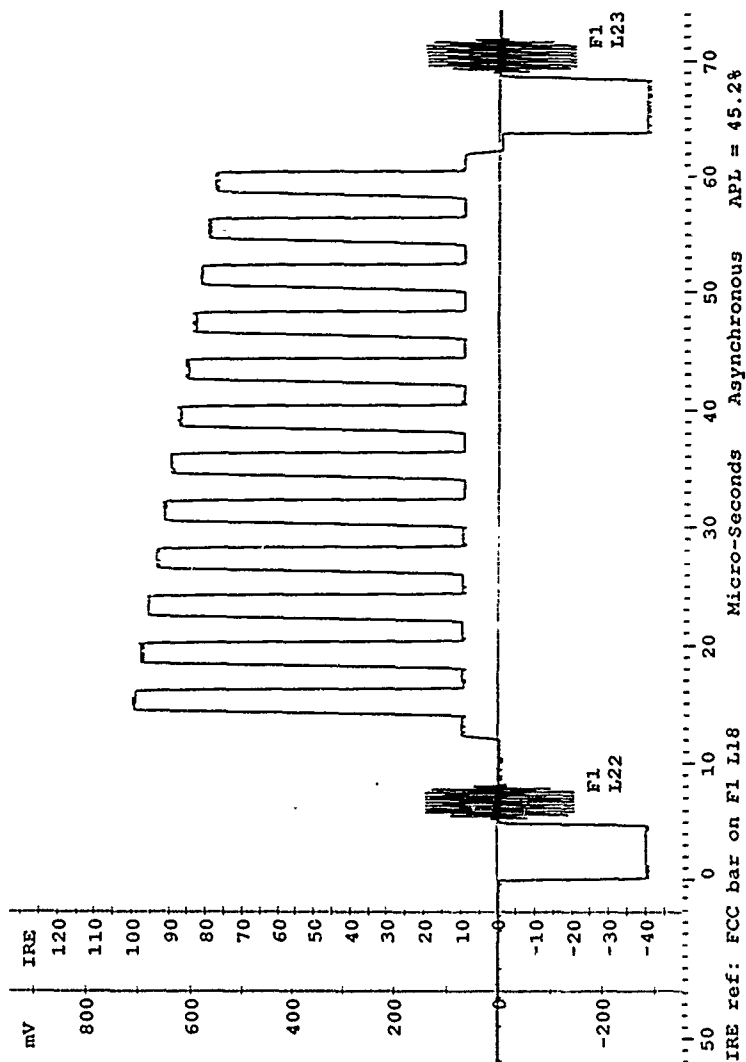


FIGURE 3-8 MULTI-STRIP - 11M BOTTOM

VM700 Video Measurement Set

Channel A Channel A

12-May-88 15:54:10

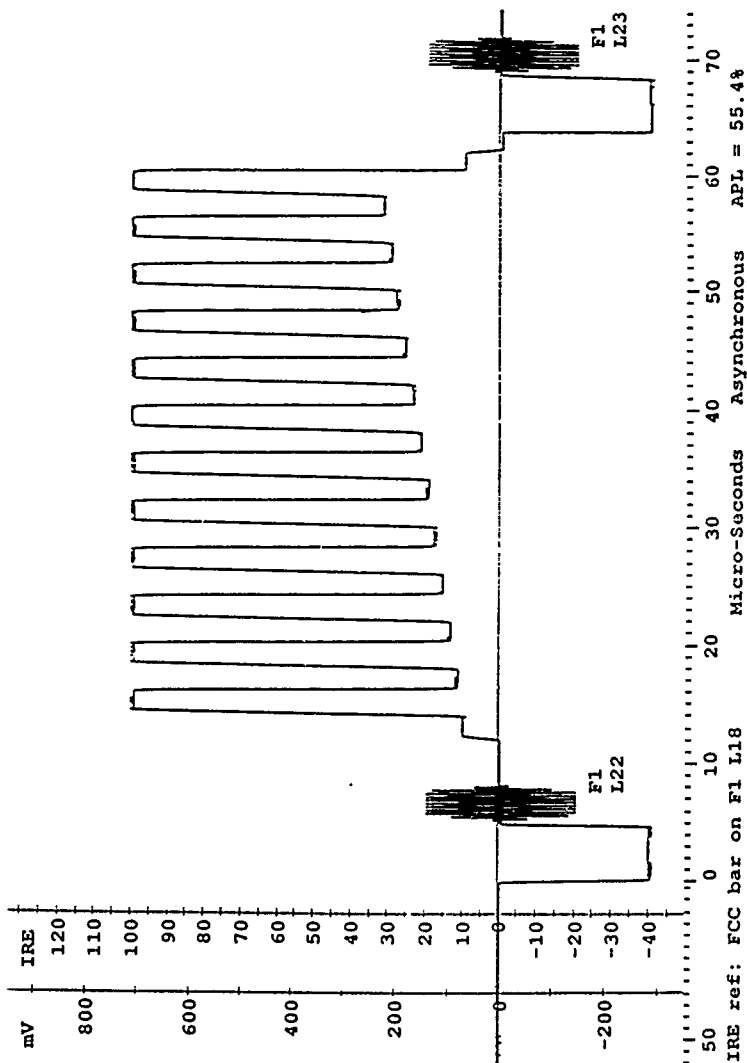


FIGURE 3-9 MULTI-STEP - 111A1 TOP

Table 3-1
CODEC Test Signals

Performance Parameter	Test Signal
Amplitude/Frequency Response	Multiburst
Insertion Gain Line Time Waveform Distortion	Line Bar (18 usec)
Short Time Waveform Distortion	2T Pulse
Chrominance-Luminance Gain Inequality Chrominance-Luminance Delay Inequality	Chrominance Pulse (Modulated 20T Pulse)
Luminance Non-Linear Distortion	Modulated 5-Riser Staircase
Differential Gain	(50% APL)
Differential Phase	
Chrominance-to-Luminance Intermodulation	
Chrominance Non-Linear Gain	3-Level Chrominance
Chrominance Non-Linear Phase	
Vector Accuracy	Color Bar Chart
Field Time Waveform Distortion Signal-to-Noise Ratio	Field Bar

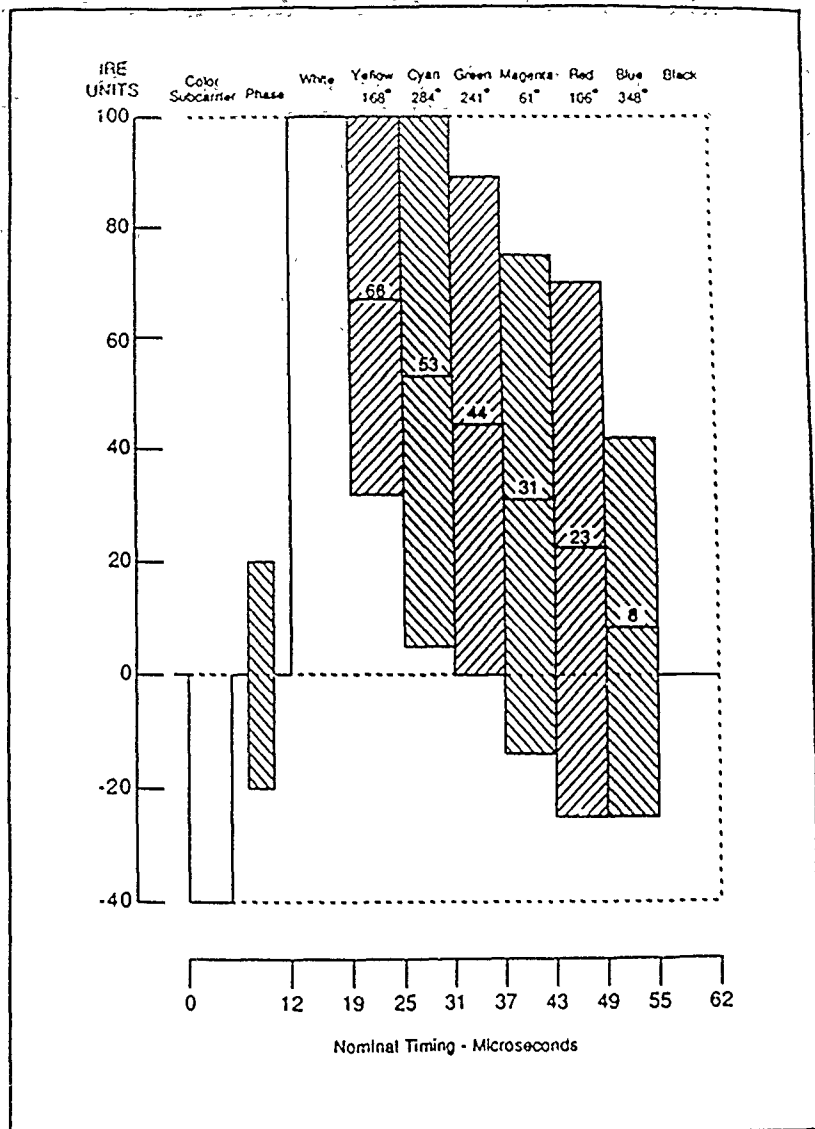


Figure 3-10 The Color Bar Chart Test Signal

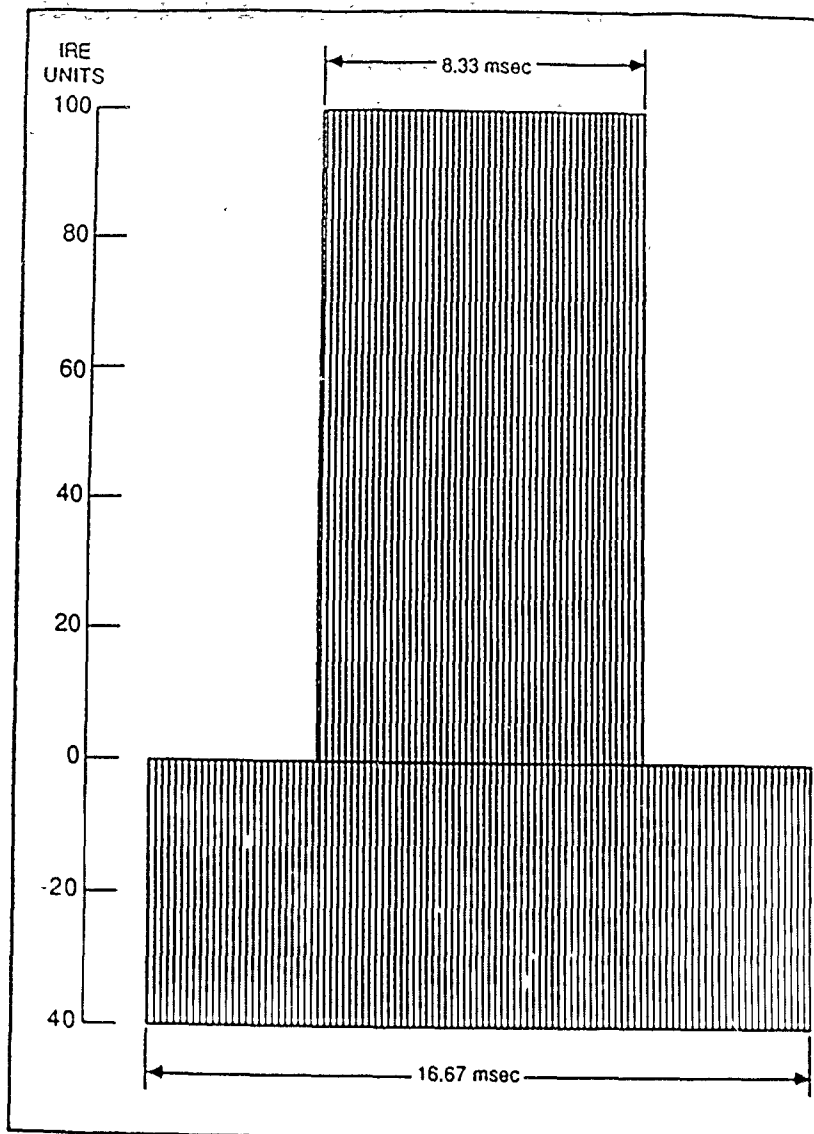


Figure 3-11 The Field Bar Test Signal

3.4 Proposed Performance Standards

A performance standard must be based on past experience and measurements. It would be useless to either impose requirements which are too difficult to meet or set limits which can be satisfied without effort and could result in a degraded picture. Measurements were made on several codecs, but they all operated at 1.544 Mbps. It is not known at this time if and how much higher still picture deterioration is to be expected at lower bit rates down to 64 Kbps. Very recent measurements on codecs operating in the 64 to 364 Kbps range show that the necessary lower sampling rate in these codecs requires narrower transmit and receive filters producing a lower overall frequency response. Correlation with subjective results will follow.

It is customary to establish not only a reference standard with an allowable deviation for each parameter but also caution and alarm limits. This originated in the broadcast industry where continuous monitoring or frequent measurement of most parameters is necessary to maintain high picture quality. Reaching a caution limit is an alert of a potential problem while exceeding an alarm limit requires immediate corrective action. At present, video teleconferencing systems are far from the sophistication of broadcast operations but this situation may change before too long, so following broadcast practices as much as possible becomes desirable.

Previously established and proposed waveform parameter standards are listed on Table 3-2. The codec standard prepared

WAVEFORM PARAMETER STANDARDS AND LIMITS

TABLE 3-2

Parameter	Meas. Units	RS-170A Standard	CODEC Standard	Caution Limits		Alarm Limits	
				Lower	Upper	Lower	Upper
Field 1	Lines	20 \pm 6	23	--	--	--	--
Vertical Blanking Width	Lines	20 \pm 6	23	--	--	22	23.1
Field 2	Lines	20 \pm 6	23	--	--	22	23.1
Horizontal Blanking Width	usec	10.9 \pm 1.2	12.1 \pm .3	11.7	12.5	11.5	12.7
Horizontal Synch Width	usec	4.7 \pm .1	4.7 \pm .2	4.4	5.0	4.2	5.2
Front Porch Width	usec	1.5 \pm .1	1.5 \pm .1	1.3	1.7	1.2	1.8
Breezeway Width	usec	.6 \pm .1	.6 \pm .1	.4	.8	.3	.9
Color Burst Width	Cycle	9	9 \pm .5	8.5	9.5	8	10
Equalizing Pulse Width	usec	2.3 \pm .1	2.3 \pm .2	2.1	2.5	2.0	2.6
Serration Width	usec	4.7 \pm .1	4.7 \pm .2	4.4	5.0	4.2	5.2
Horizontal Synch Rise Time	nsec	140 \pm 20	140 \pm 30	100	180	80	200
Horizontal Synch Fall Time	nsec	140 \pm 20	140 \pm 30	100	180	80	200
White Bar Amplitude	IRE	100 \pm 2	100 \pm 2	94	106	90	110
Synch Amplitude	IRE	40 \pm 2	40 \pm 2	37	43	35	45
Burst Amplitude	IRE	40 \pm 2	40 \pm 2	37	43	35	45

by Delta Information Systems for the Defense Communications Agency (DCA) is essentially identical with the values specified in RS-170A. The only major deviation is in the blanking widths which are matched to one specific codec design for which they must be accurately maintained. Therefore, the limits are set very tight. Another codec is likely to require different numbers, so a universal standard is difficult to establish. The only possibility would be to impose maximum values to restrict the permissible amount of picture area clipping. A survey of all available codecs is needed to arrive at such a practical limit. Different values of vertical blanking width may apply to alternate fields.

Though most caution limits are set somewhat beyond the proposed standard, they are well within the range of completely acceptable performance. Exceeding the caution limit may result in a barely noticeable deterioration. However, exceeding the alarm limit is likely to cause sufficient picture degradation to require remedial action. The time and type of such action depends on the operational requirements of the system.

Table 3-3 lists the proposed standards, caution and alarm limits for the applicable performance parameters taken from RS-250B and NTC Report No. 7. A composite of the values of these two documents is included as reference standard. These values are generally close to the end-to-end requirements of RS-250B but not necessarily as lenient as the values of NTC-7. The remarks of the preceding paragraph regarding the caution and alarm limits

PERFORMANCE PARAMETER LIMITS

TABLE 3-3

PARAMETER	MEAS. UNITS	REFERENCE STANDARD	CODEC STANDARD	CAUTION LIMIT	ALARM LIMIT
Amplitude/Frequency Response	dB	+6/-1to 4.2 KHz	-1 @ 1MHz -1.6 @ 2MHz -3 @ 2.7MHz	-1 @ 1MHz -2 @ 2MHz -4 @ 7MHz	-1.5@1MHz -2.5@2MHz -5@2.7MHz
Signal-to-Noise Ratio	dB	54	52	50	47
Insertion Gain	dB	+5	- 5	+6	1.0
Field Time Waveform Distortion	IRE	3	3	3	5
Line Time Waveform Distortion	IRE	3	3	3	5
Short Time Waveform Distortion	% FAC	3	4	4	6
Chrominance-Luminance Gain Inequality	%	7	±10	±10	±15
Chrominance-Luminance Delay Inequality	nSec	60	+75	+100	±150
Luminance Non-Linear Distortion	IRE	10	10	10	20
Differential Gain	%	10	10	10	15
Differential Phase	Degree	5	5	6	10
Chrominance-to-Luminance Intermodulation	IRE	4	4	4	6
Chrominance Non-Linear Gain	IRE	20± 2 80± 8	20±2 80±8	20±2 80±8	20±3 80±12
Chrominance Non-Linear Phase	Degree	5	5	5	7
SC-H Phasing	Degree	±50	±60	±75	±100
Vector Accuracy	Amplitude Error	%	—	±10	±20
	Phase Error	Degree	—	± 5	±10
	Chroma/Lum Ratio	%	—	±10	±20

are equally applicable to this set of parameters.

Basing the codec standard on a tighter reference standard would be counterproductive in view of the inherent performance limits of any teleconferencing codec operating in the 64 to 1544 Kbps range. It might result in the rejection of a codec for not meeting a performance specification which is not relevant for the application intended by NCS or other users as shown by correlation of objective and subjective tests.

The only parameter significantly different from broadcast requirements is amplitude/frequency response which is inherently limited in every codec. Experience has shown that judicious filtering of the output video signal can reduce minor disturbances caused by the compression algorithm and thus produce a more pleasing picture. Measuring the codec response is best done with a modified multiburst signal. Review of the requirements and performance data has shown that the frequency packets should be .5, 1.0, 2.0, 2.35, 2.7 and 3.0 MHz, as shown on Figures 3-1 to 3-3. The other modified test signal is the chrominance pulse which is recommended with a width of 20T to avoid any potential limitations caused by chrominance circuit frequency response.

The frequency response standards and the choice of the frequency packets are based on measurements on existing 1.544 Mbps codecs. It is likely that the response of low bit rate codecs is more limited in which case different values may be more desirable. The same could be true with some other parameters,

with some other parameters, though it is expected that most of the listed values will be realistic for all codecs. On the other hand, some codecs may have a much better frequency response in a freeze-frame mode. If necessary, it is easy to change the PROMs in the 1910 signal generator back to normal, which will result in frequency packets of .5, 1, 2, 3, 3.58, and 4.2 MHz.

Two parameters not used in RS-250B and NTC-7 have been added. They are Subcarrier to Horizontal Sync (SC-H) Phasing and Vector Accuracy. SC-H Phasing is specified conditionally in RS-170A. It does not affect the directly viewed picture but is important for video tape recording which is likely to be used extensively in future teleconferencing systems. Vector Accuracy provides a convenient overview of color performance and is very useful for a "quick look", ahead of a more detailed analysis of the received video signal. There is no reference standard for Vector Accuracy, so the proposed codec standard is based on practical experience.

SECTION 4 - MOVING PICTURE MEASUREMENTS

4.1 Background

Past experience has shown the motion rendition of a video teleconferencing system to be the most important parameter which affects overall picture quality and user acceptance. The only practical method of codec evaluation so far has been strictly subjective which is inaccurate, complex, time consuming and costly. This points out the necessity of developing a methodology to objectively test codec motion rendition capability. Without this it is impossible to establish a system performance standard.

The initial small steps towards this objective were taken by DIS following the subjective comparative evaluation of four codecs operating at 1.544 Mbps. The test signal portion of the test tape contained two sequences intended for the purpose of motion evaluation with pixel changes produced by switching between either the "white window" signal and a black field or yellow and blue fields.

Switching between colors yielded no usable result but the white window/black field test showed promise. Some low level residue of the white area was clearly visible for several frames after switching. In addition, the codec with the subjectively poorest performance was even unable to follow a normal switch between two average pictures.

It is important to be able to correlate objective and

subjective codec evaluations to confirm the validity of any new objective method. The objective test should be performed with a carefully controlled video signal which can be observed simultaneously on a waveform monitor and a picture monitor. Actual measurements can be made only on the waveform monitor but the picture monitor helps in interpreting the observed signal waveform. It would be a mistake to try to subjectively evaluate the picture on the monitor since it shows an artificial test sequence with no relation to a typical live scene. Subjective tests on several codecs at various bit rates with different conventional motion scenes will establish a threshold of acceptability which can then be correlated with an objective test score. The objective test material must cover a range of conditions wide enough to produce both acceptable and unacceptable performance of codecs of different quality over the full range of commonly used bit rates.

The motion performance of a codec is determined by its capability to faithfully reproduce pixel changes between consecutive frames. Such changes can be introduced either by switching between different pictures (scene cut) or by changes within the picture. These two types of pixel changes generally produce rather different visually noticeable picture deteriorations. Experience with subjective tests has indicated that similar ratings of codec performance are obtained by observing its response to either a scene cut or motion within the picture. However, largely due to the differences in picture

degradations it has so far been difficult to establish reliable correlation between the two types of test material. It has therefore been found necessary to investigate both scene cuts between accurately defined pictures and controlled artificial motion as candidate objective motion test signals. A detailed description of the considerations for the design of the motion test sequences is given in the two subsequent paragraphs.

A practical way of producing motion test patterns is by means of a computerized generator of animated pictures. The TV production studio which has performed services for DIS over several years owns the Cubicomp equipment which has ample capability and flexibility to meet all requirements. Each picture is first composed, then reviewed, and finally recorded frame-by-frame on 1" video tape.

4.2 Scene Cut Patterns

A scene cut with a large and accurately defined pixel change can be implemented by alternating between a white on black pattern and an all-black picture. Previous tests with the conventional white window signal indicated that this method was feasible but a pattern with much more detail was needed to sufficiently stress the capability of the codec. A logical expansion of the window would be a checkerboard of varying sizes, but circles were considered preferable because curved contours are more challenging for codecs using transform coding and because they provide two independent variables: center spacing

and radius. Furthermore, in order to minimize repetition of possible blocks subjected transform coding, the axis connecting the centers of the circles is tilted 25 degrees to the horizontal. Three circle spacings are employed, each with two different radius values which produce about 60% and 30% pixel changes. The smaller spacings result in more variety in the pattern and thus in increased challenge to the codec algorithm. However, both circle spacings and radius are always kept large enough so that resolution limitations do not become a determining factor. A typical pattern is shown on Figure 4-1.

The speed of consecutive scene cuts covers a wide range. The lowest switching rate allows enough settling time for probably any codec to reach a stable picture. The fastest rates are expected to be challenging to any codec and produce some deterioration in all cases. The duration of each scene is 15 seconds.

The circle pattern was produced on the Cubicomp Special Effects Equipment which was easily programmed for this purpose. The values of circle spacing, radius and switching rate are listed on Table 4-1. Actual dimensions would be meaningless because they depend on monitor screen size and edge masking imposed by both codec and monitor; therefore, they are given as percentages of the full picture width with standard horizontal blanking. The switching rate is specified as the number of frames between switches which is the term required by the Cubicomp Equipment. The combinations of circle dimensions and

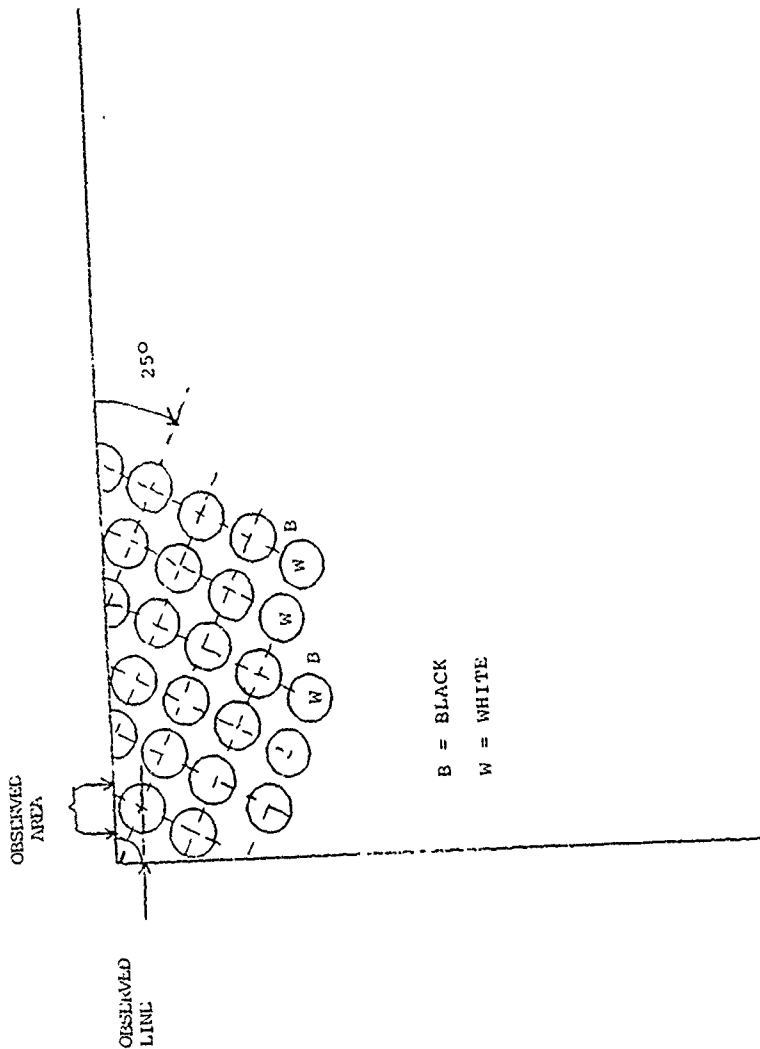


FIGURE 4-1 CIRCL PATTERN

CIRCLE SPACING (%)		7			4			2.25	
CIRCLE RADIUS (%)		3.25	2.25	1.75	1.25	1	.75		
120		A-1	-	A-13	-	A-25	-		
60		A-2	-	A-14	-	A-26	-		
30		A-3	-	A-15	-	A-27	-		
15		A-4	-	A-16	-	A-28	-		
8		A-5	A-9	A-17	A-21	A-29	A-33		
4		A-6	A-10	A-18	A-22	A-30	A-34		
2		A-7	A-11	A-19	A-23	A-31	A-35		
1		A-8	A-12	A-20	A-24	A-32	A-36		

SWITCHING
RATE
(FRAMES)

TABLE 4-1
SCENE CUT PATTERNS

switching rates which are implemented on the tape are indicated by insertion of a scene number (such as A-1) on the table.

4.3 Rotating Wheel Patterns

Artificial motion can be implemented in various ways. The simplest form is linear, such as a horizontal, vertical or diagonal wipe, which produces motion in just one direction. Rotation or zooming results in motion which varies in direction all over the picture. This is more desirable because linear motion may give an undue advantage to some codecs using motion vectors for interframe coding. Though in actual use zooming is much more common than rotation, it does not lend itself readily to objective evaluation on a waveform monitor. However, a rotating pattern can be designed to contain areas in which motion produces fully predictable waveforms which can be analyzed on a waveform monitor. Any motion degradation will then manifest itself as easily recognizable and measurable changes in this waveform display.

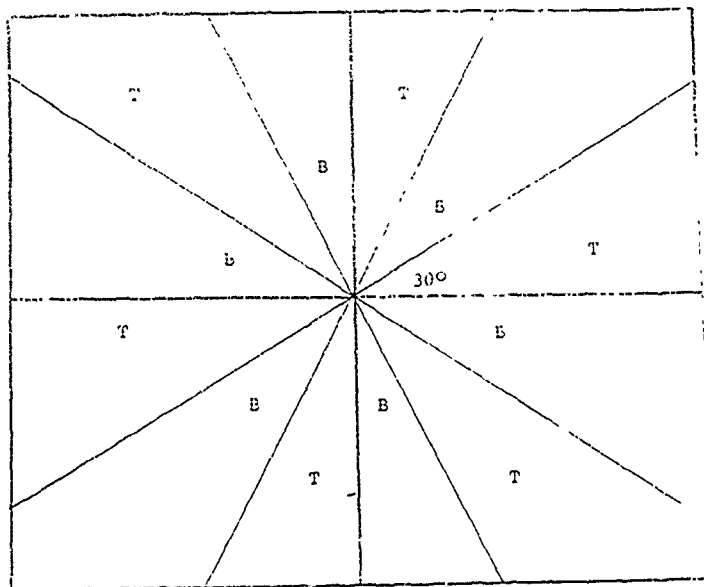
Motion degradation is subjectively equally noticeable in monochrome and in color. Objective evaluation is more practical with a monochrome signal because the presence of color subcarrier makes measurements more difficult, and use of a luminance filter is not always practical. However, the presence of color in the picture increases the overall challenge on the codec algorithm and thus is likely to at least indirectly affect motion performance. Therefore, a rotating wheel pattern containing both

monochrome and color portions is needed. Different widths of the wheel spokes result in varying numbers of transitions per revolution to be reproduced. Combined with a wide range of rotation speeds this will encompass the performance limits of all codecs to be evaluated.

The rotating wheel patterns were also produced on the Cubicomp Special Effects Equipment which determined some of their features. One of the patterns that could be produced without excessive complexity is shown on Figure 4-2. It consists of two portions, a fixed background with a white polygon in the center and segments of the six standard bar chart colors around the edges, and rotating black spokes covering the whole picture. This results in transitions both between black and white and between black and color. For compatibility with the Cubicomp equipment, the speeds of rotation had to be specified as the number of frames per full revolution, with the constraint that the increments between frames must be integral numbers of degrees.

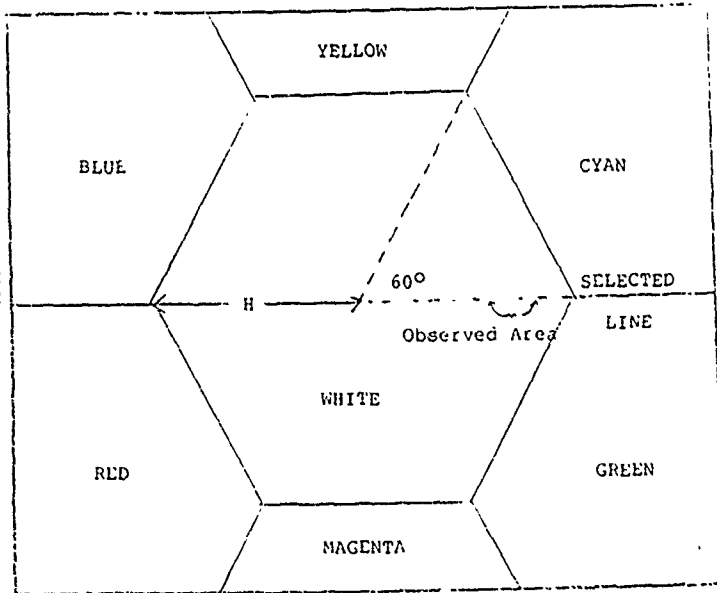
Wheel patterns with three different spoke widths were implemented. The properties of the various scenes produced by different rotation speeds are summarized in Table 4-2. Several interrelated parameters are given to describe the rotation speed of the patterns because, depending on the purpose of a specific analysis, different points of view have to be taken. It is first given in frames per revolution to satisfy the requirements of the production equipment. In conventional terms, degrees per second

B = BLACK
T = TRANSPARENT



ROTATING OVERLAY

H = APPROX.
40% OF
VISIBLE
PICTURE
HEIGHT



BACKGROUND

FIGURE 4-2 30 DEGREE WHEEL PATTERN

SPOKE WIDTH 30°

SCENE NO.	B-1	B-2	B-3	B-4	B-5	B-6	B-7	B-8	B-9
FRAMES/REVOL.	540	360	240	180	144	120	90	72	60
DEGREES/SEC.	20	30	45	60	75	90	120	150	180
EDGES/SEC.	.67	1	1.5	2	2.5	3	4	5	6
FRAMES/SPOKE	45	30	20	15	12	10	7.5	6	5
% PIXEL CHANGE PER FRAME	2.2	3.3	5	6.7	8.3	10	13.3	16.7	20
% BLOCK CHANGE PER FRAME	18			22			34		43

SPOKE WIDTH 18°

SCENE NO.	B-1	B-2	B-3	B-4	B-5	B-6	B-7	B-8
FRAMES/REVOL.	720	540	360	240	180	144	120	90
DEGREES/SEC.	15	20	30	45	60	75	90	120
EDGES/SEC.	.83	1.1	1.7	2.5	3.3	4.2	5	6.7
FRAMES/SPOKE	36	27	18	12	9	7.2	6	4.5
% PIXEL CHANGE PER FRAME	2.8	3.7	5.6	8.3	11.1	13.9	16.7	22.2
% BLOCK CHANGE PER FRAME	31				42			57

TABLE 4-2 SHEET 1
ROTATING WHEEL PARAMETERS

SPOKE WIDTH 10°

SCENE NO.	B-18	B-19	B-20	B-21	B-22	B-23
FRAMES/REVOL.	720	540	360	240	180	144
DEGREES/SEC.	15	20	30	45	60	75
EDGES/SEC.	1.5	2	3	4.5	6	7.5
FRAMES/SPOKE	20	15	10	6.7	5	4
% PIXEL CHANGE PER FRAME	5	6.7	10	15	20	25
% BLOCK CHANGE PER FRAME	50	54			70	75

TABLE 4-2 SHEET 2
ROTATING WHEEL PARAMETERS

is much more descriptive. When considering picture transitions, the number of edges per second is useful. In low bit rate codecs which may employ a large amount of frame repetition, the number of frames per spoke helps in understanding some phenomena and unusual results. The number of pixel changes between frames gives a good measure of the load on the codec algorithm.

All these parameters are easily computed accurately. However, it is understood that in the case of transform coding the percentages of change of encoded blocks are the ultimate criteria of the stress on the codec algorithm but these numbers depend on too many variables such as the exact location of each block in relation to the picture pattern to be readily computed. Assuming a picture structure of 36 x 44 blocks, an approximate graphic method was used to estimate this number for some of the spoke width/rotation speed combinations. The results show that, while the numbers of pixel changes overlap over a wide range of rotation speeds for the three spoke widths in use, the block change numbers are not only much higher but also have a much smaller range of overlap. The same number of pixel changes produces a higher number of block changes with narrower spokes. This is important when determining the required range of spoke widths and rotation speeds for testing codecs of different quality operating at different bit rates.

The constraints of the Cubicomp equipment called for some special arrangements to produce all desired rotation speeds. At low speeds with more than 360 frames per revolution it is

necessary to repeat every or every other frame. Some intermediate speeds are produced by advancing alternate frames by different numbers of degrees. These necessary slight discontinuities are essentially invisible on the resulting tape and do not affect the analysis of the resulting waveform patterns. The motion arrangement for alternate frames to produce all implemented rotation speeds is shown on Table 4-3.

4.4 Test Tape Preparation

The test tape containing all the scenes listed on Tables 4-1 and 4-2 was prepared in 1" format as required by the Cubicomp equipment and subsequently reduced to 3/4" format for convenience in use. Since the purpose of this interim version of the tape was to establish the validity of this motion test method and to develop a measurement method, no editing was performed. Though the durations were planned to be 15 seconds for "A" and 30 seconds for "B" scenes, some of these numbers were shortened at high switching rates or speeds for the purpose of convenience and economy without sacrificing usefulness. Each scene was identified informally at the start with its "A" or "B" number without any further description of its content.

A final motion test tape will be prepared as part of a different task after tests with the existing interim tape have shown the usefulness of the various scenes. It is expected that this final version may contain fewer and/or shorter scenes, though, if deemed necessary, new scenes with the same basic

FRAMES/ REVOL.	FRAME #1		FRAME #2	
	FRAMES HELD	DEGREES ADVANCED	FRAMES HELD	DEGREES ADVANCED
720	2	1	2	1
540	2	1	1	1
360	1	1	1	1
240	1	2	1	1
180	1	2	1	2
144	1	3	1	2
120	1	3	1	3
90	1	4	1	4
72	1	5	1	5
60	1	6	1	6

TABLE 4-3
MOTION ARRANGEMENTS

patterns but different speeds can be added without undue effort.

4.5 Measurement Implementation

Instrumentation for the evaluation of processed motion test patterns is rather different from conventional video measurements. Observation on a waveform monitor at either line or field rate yields little or no useful information. The scene cut patterns show that at high switching rates residual signals make it impossible to reach the full white and black levels, so that the video peak-to-peak amplitude is reduced, but this feature does not lend itself to numerical evaluation. The rotating wheel conventional waveform display is too complex for any meaningful evaluation.

To avoid overlap of different portions of the picture on the waveform monitor screen, it is necessary to view a suitable selected single line. This line selection feature is available on most up-to-date waveform monitors. It allows observation of a single line which can be highlighted on an associate picture monitor, and frequently the line number is displayed in a corner of the waveform display. Sweep expansion and positioning make it possible to examine a portion of the selected line in greater detail.

Observation of a changing picture waveform at all but very low speeds or switching rates is impossible. Therefore it is necessary to examine the waveform on a frame-by-frame basis. This is presently possible only with a top grade 1" video tape

recorder with still picture capability, manual tape advance and frame counter. Though a still picture can be viewed also on many 3/4" tape recorders, the vertical interval waveform in this mode is presently not compatible with line selection, therefore 3/4" tape has so far not been usable. Thus, even though 3/4" tape is adequate for processing through the codec, it had to be transcribed to 1" format for motion performance measurements.

The test setup is shown on Figure 4-3. The Sony BVH-2500 and the Tektronix 1730 have all necessary features. There are no special requirements on the monitor. Having all three equipments in the same rack results in a convenient setup for checking the effect of each frame advance and correlating the waveform with the displayed picture.

Figures 4-1 and 4-2 show the approximate locations of the observed areas for both patterns. In the case of the scene cut, a white circle near the upper left corner of the picture was used with the selected line near the center of the circle. A line in the top half of the picture is desirable because at times switching causes some added disturbances in the bottom half. The measurements on the rotating wheel pattern were performed on the easily identified center line in the area where the picture change is between black and white. The choice of the right side of the picture was arbitrary.

Even before any numerical evaluation, some limiting effects of codec design features became obvious. The main factor is frame rate subsampling or frame repetition. Since most codecs

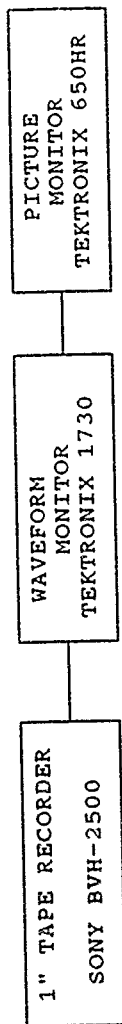


FIGURE 4-3
TEST SETUP FOR MOTION PERFORMANCE EVALUATION

transmit as a maximum only every other frame, the fastest scene cut switching rate - once every frame - becomes useless because it results in a still picture. Low bit rate codecs transmit even fewer frames and therefore have to resort to higher frame repetition numbers. If this number is higher than the number of frames per spoke (see Table 4-2) one spoke may at times be skipped completely, while the duration of the two adjacent spokes appears to be unduly extended, which produces an invalid result. In a fair number of higher speed cases a stroboscopic effect caused by interaction of pattern motion and frame repetition becomes obvious. In such cases, numerical results must be carefully scrutinized and checked for their validity. Interaction of block coding, frame subsampling and high speed rotation often produces artifacts which are not readily recognizable in the waveform pattern and therefore make the measured results difficult to interpret.

SECTION 5 - MOTION TEST METHODOLOGY ANALYSIS

5.1 Data Collection

Since the purpose of this program is the establishment and verification of a new methodology for objective evaluation of motion performance, no attempt was made to make comparative tests of codecs. However, it is necessary to gather data from several codecs operating over the range of conventional bit rates in order to find the dependencies of the measured pattern results on test parameters and the features of the codecs under test.

The test tape described in the preceding section was processed through three codecs at bit rates ranging from 64 Kbps to 1.544 Mbps. The results were recorded on 3/4" tape and first examined visually to obtain a preliminary review. Next the tapes containing the data most likely to be significant for verification of the test methodology were transcribed to 1" format and examined using the test set-up shown on Figure 4-3. The amount of available data was much larger than deemed necessary. Therefore a judicious selection of test patterns, switching rates and rotation speeds processed through the available codecs at various bit rates was made to determine the combinations of these parameters that yield the most meaningful results.

The waveform of each frame ideally shows a flat white or black level in the area selected for examination. The conventional measurement of this level is in IRE units.

Secondary effects, caused for instance by interaction of coding blocks with the pattern and similar artifacts, which are clearly visible after passage of the signal through the codec produce some disturbances which result in potential errors of about ± 3 IRE units. The frame-by-frame analysis ideally produces a full amplitude square pattern of video level vs. frame number. The deviations of the actual processed pattern from this ideal are a measure of the motion distortion produced by the codec.

The number of frames to be examined varies widely depending on the rate of occurrence of white to black and black to white transitions. In order to achieve confidence in the result, several cycles of these changes must be examined. The minimum number was 4 but up to about 10 cycles were used at higher speeds. This means up to about 400 frames at low speeds, while at high speeds 40 to 30 frames are sufficient to show a typical repetitive pattern. Since at all but the highest speeds the video amplitude remains constant for several frames, it is not necessary to record the level of every frame. Noting the first and last frame numbers of a series of constant amplitude frames in addition to all level changes is sufficient and simplifies both the data taking and evaluation efforts. For numerical evaluation, all data points may be assumed to be connected by straight lines.

5.2 Numerical Evaluation

A large number of data were taken on the tapes processed

through the three available codecs at bit rates of 64 and 384 Kbps and 1.544 Mbps. Both scene cut (A) and rotating wheel (B) test sequences were used. Figure 5-1 shows the very simple sheet used to record the test data. The header of the sheet gives the pertinent data to identify each test, an additional single letter (C, L, P) was used to designate the codec under test.

The amplitude extremes at low speeds were in the vicinity of 10 and 100. Minor deviations were due mainly to small changes in setup and do not affect the validity of the data. As expected, at higher speeds the excursions between the amplitude extremes become smaller and many intermediate levels with varying durations appear. This is due to the inability of the codec to follow fast changes in amplitude and thus is a direct indication of its motion performance capability.

The method of numerical evaluation of the results has to be chosen very carefully. At first, curves for some tests were drawn by hand and the difference between the estimated arithmetic averages of the extremes of each curve was determined. The irregularity of many of the curves made this method rather awkward and the accuracy of the results was questionable. The RMS value of the excursions from the mean level of each curve appears to be preferable because it emphasizes the important maximum excursions and their duration. This process was implemented with a computer program which provided a printout of the curve and the RMS value for each test sequence. A typical

CODEC OBJECTIVE MOTION TEST DATA

BIT RATE

TEST SEQUENCE: LINE NO: LOCATION: APPROX. SEC FROM LINE START

FRAME	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
AMPL.																														
FRAME	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
AMPL.																														
FRAME	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
AMPL.																														
FRAME	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
AMPL.																														
FRAME	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
AMPL.																														
FRAME	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180
AMPL.																														
FRAME	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210
AMPL.																														
FRAME	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240
AMPL.																														
FRAME	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270
AMPL.																														
FRAME	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300
AMPL.																														

FIGURE 5-1 TEST DATA RECORDING FORM

printout is shown on Figure 5-2.

Each printout presents the processed result of one full set of measured data with one specific test pattern but provides merely one point on a curve describing codec motion performance. It is necessary to define a common parameter for these curves. Such a parameter is temporal frequency which can be expressed in cycles per second where each cycle is produced by one white-black-white change. Plotting RMS amplitude vs. temporal frequency gives a temporal response curve. This concept is equally applicable to scene cut and rotating wheel sequences. The temporal frequencies produced by each test sequence are listed on Table 5-1. The numbers are readily derived from the data on Tables 4-1 and 4-2. It must be noted that the temporal frequency is not the only parameter that determines the result. The configuration of the test pattern is of equal importance. The same temporal frequency produces different RMS amplitudes when different test patterns are used.

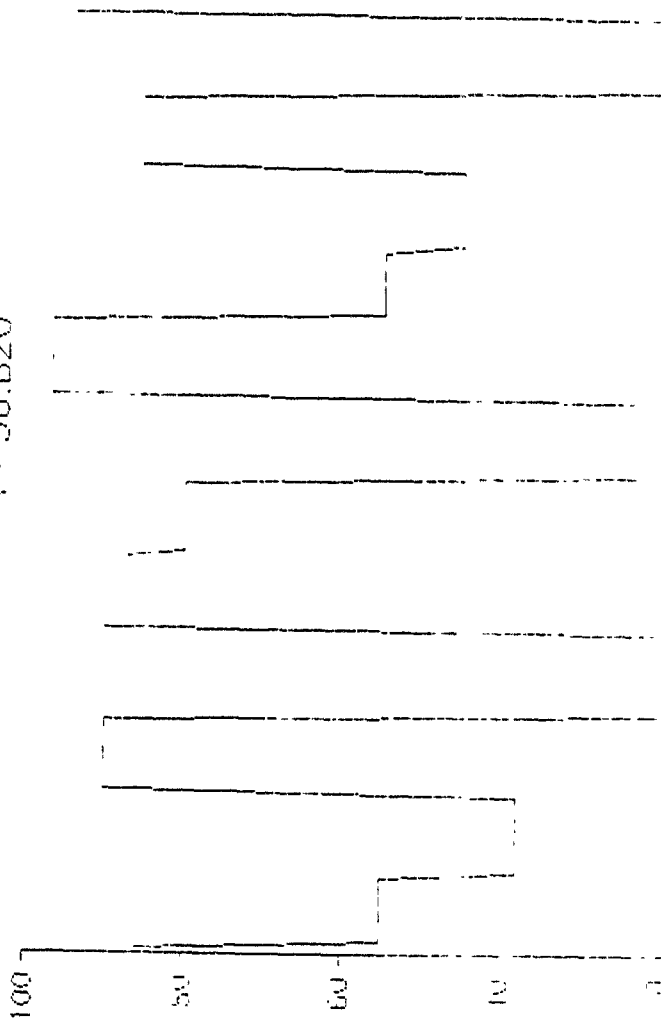
5.3 Results

The test results are summarized on graphs of RMS amplitude vs. temporal frequency for a specific bit rate and codec, using one test pattern over the available range of either switching rates or rotation speeds. A number of representative graphs are shown on Figures 5-3 to 5-16. At least 2 graphs are on each figure to illustrate the effect of changing a single parameter. It should be noted that the data points were connected by

SEQUENCE

C-56.B20

24.678000



FRAME

FIGURE 5-2: WAVEFORM & COMPUTED RMS VALUE

SEQUENCE NO.	TEMPORAL FREQUENCY
A-1, A-13, A-25	.125
A-2, A-14, A-26	.25
A-3, A-15, A-27	.5
A-4, A-16, A-28	1.0
A-5, A-9, A-17, A-21, A-29, A-33	1.875
A-6, A-10, A-18, A-22, A-30, A-34	3.75
A-7, A-11, A-19, A-23, A-31, A-35	7.5

SEQU. NO.	TEMP. FREQU	SEQU. NO.	TEMP. FREQU	SEQU. NO.	TEMP. FREQU.
B-1	.33	B-10	.42	B-18	.75
B-2	.5	B-11	.55	B-19	1.0
B-3	.75	B-12	.85	B-20	1.5
B-4	1.0	B-13	1.25	B-21	2.25
B-5	1.25	B-14	1.67	B-22	3.0
B-6	1.5	B-15	2.1	B-23	3.75
B-7	2.0	B-16	2.5		
B-8	2.5	B-17	3.3		
B-9	3.0				

TABLE 5-1 TEMPORAL FREQUENCIES

straight lines; no attempt was made to draw a smooth curve. This was done because the purpose of these tests is to verify the test methodology and not to evaluate codec performance. All test points must be considered equally valid and any deviations should not be routinely smoothed out but investigated, explained and used as a guide for the development of a more refined methodology.

Figures 5-3 to 5-11 contain temporal frequency responses produced with the three rotating wheel patterns. Specifically, Figures 5-3 to 5-5 show the dependence on the selected pattern. The curves include data for all three tested codecs and bit rates. They clearly demonstrate that, as expected, the narrower spokes are more demanding on the codec at the same temporal frequency. Therefore, the spoke width should be selected to match the motion capability of the codec to produce a curve with a significant steady slope which can be easily interpreted.

Figures 5-6 to 5-8 show the effect of different bit rates when using one codec with one spoke width pattern. The response deteriorates at lower bit rates which again is logical and according to expectation. Finally, the differences between codecs is demonstrated in Figures 5-9 to 5-11. Codec L has an algorithm optimized to operate at high rates while codec C is designed for low rates. Thus 384 Kbps is the lowest possible operating rate for codec L and the highest rate for codec C. Therefore, it is expected and has been demonstrated by subjective observation that at 384 Kbps the motion performance of codec C is

superior. The temporal response curves corroborate that fact at different spoke widths as shown on Figures 5-9 and 5-10. Both codecs P and C operate at 56 Kbps and subjective observers generally prefer codec P. Again, this fact is shown in the curves of Figure 5-11.

Figures 5-12 to 5-16 contain similar curves for the scene cut patterns. Figures 5-12 and 5-13 show the effect of different circle sizes on two codecs at two bit rates. As expected, the smaller dots are more demanding and produce a lower response at a lower bit rate. Figure 5-16 shows the difference in response between codecs L and C which is less pronounced when using the switched circle pattern rather than a rotating wheel pattern.

Some of the curves contain inconsistencies. Quite often the response does not show the expected smooth drop with increasing temporal frequency. Small deviations are due to the inevitable inaccuracies of the level readings on the waveform monitor and some artifacts in the picture. A few curves have complete reversals at high temporal frequencies and low bit rates. This effect is due to "temporal aliasing" which, as previously mentioned, is caused by the frame repetition in the codec. Such portions of a response curve obviously are completely invalid and must be disregarded. With enough experience in this methodology, it should be possible to limit the applied test patterns and temporal frequencies so that this spurious effect will not occur.

The various curves have their highest point generally in the vicinity of 40. The highest theoretically possible value is

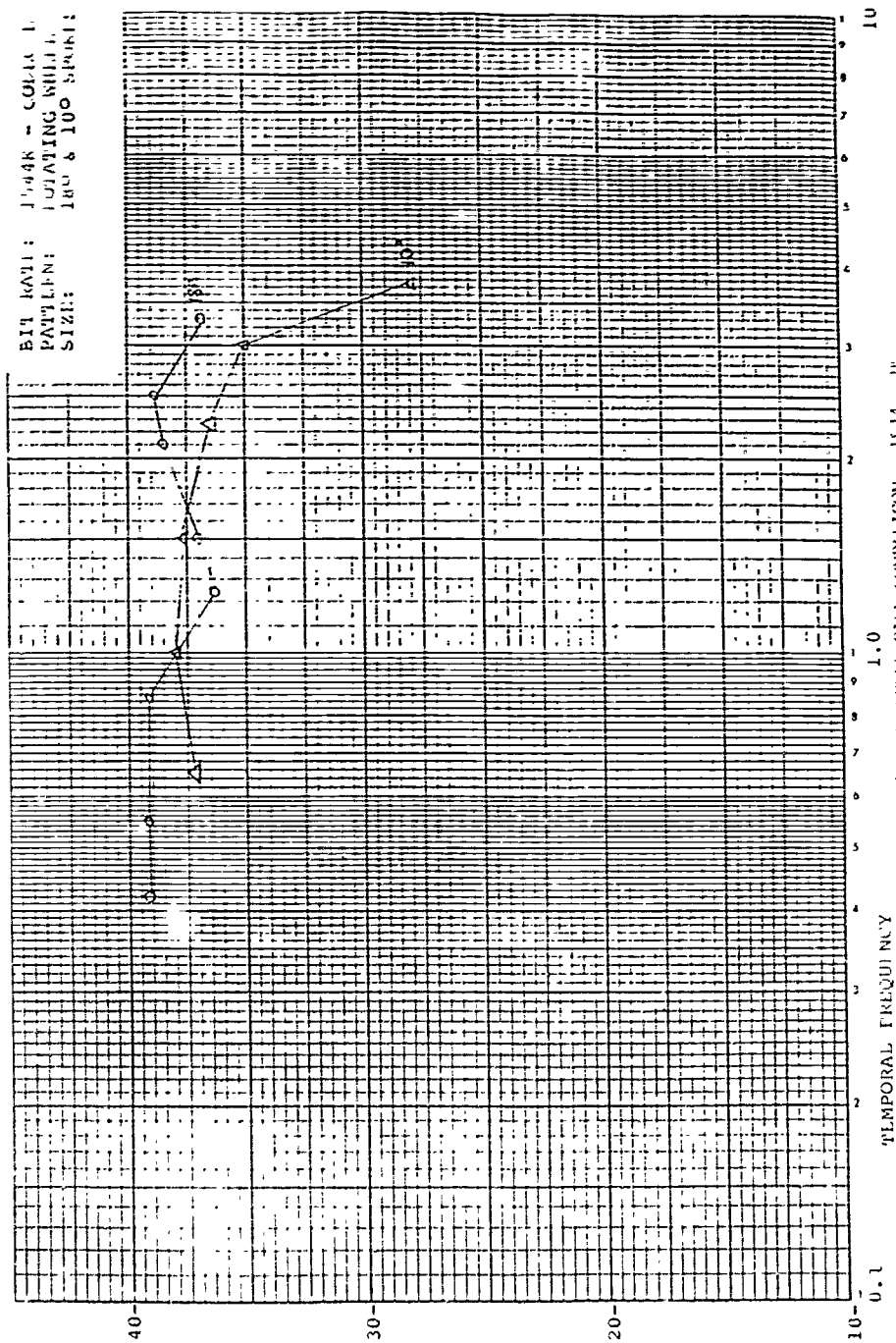
easily computed since the RMS value of a square waveform is one half of the peak-to-peak amplitude. A review of the test data sheets shows that at low temporal frequencies the black level is generally near 10 and the white level near 100 IRE Units. This produces a theoretical maximum RMS value of 45 which is in good agreement with the plotted results.

Both scene cut and rotating wheel patterns have been shown to produce useful data and thus present a valid methodology. However, the results obtained from the rotating wheel are more consistent and less subject to disturbances. Another important factor is that a scene cut produces the pixel change in an easily predictable fashion. It would therefore be possible to design a codec with an algorithm producing excellent response to a scene cut but poor actual motion performance. Testing such a codec with scene cut patterns would give completely misleading results.

Based on these considerations, it is recommended that initially only rotating wheel patterns be pursued further and incorporated in the first version of the objective motion test tape to be prepared under another task. However, the ability to respond to a scene cut is an important performance feature of a codec which should and will be investigated in addition to the temporal frequency response.

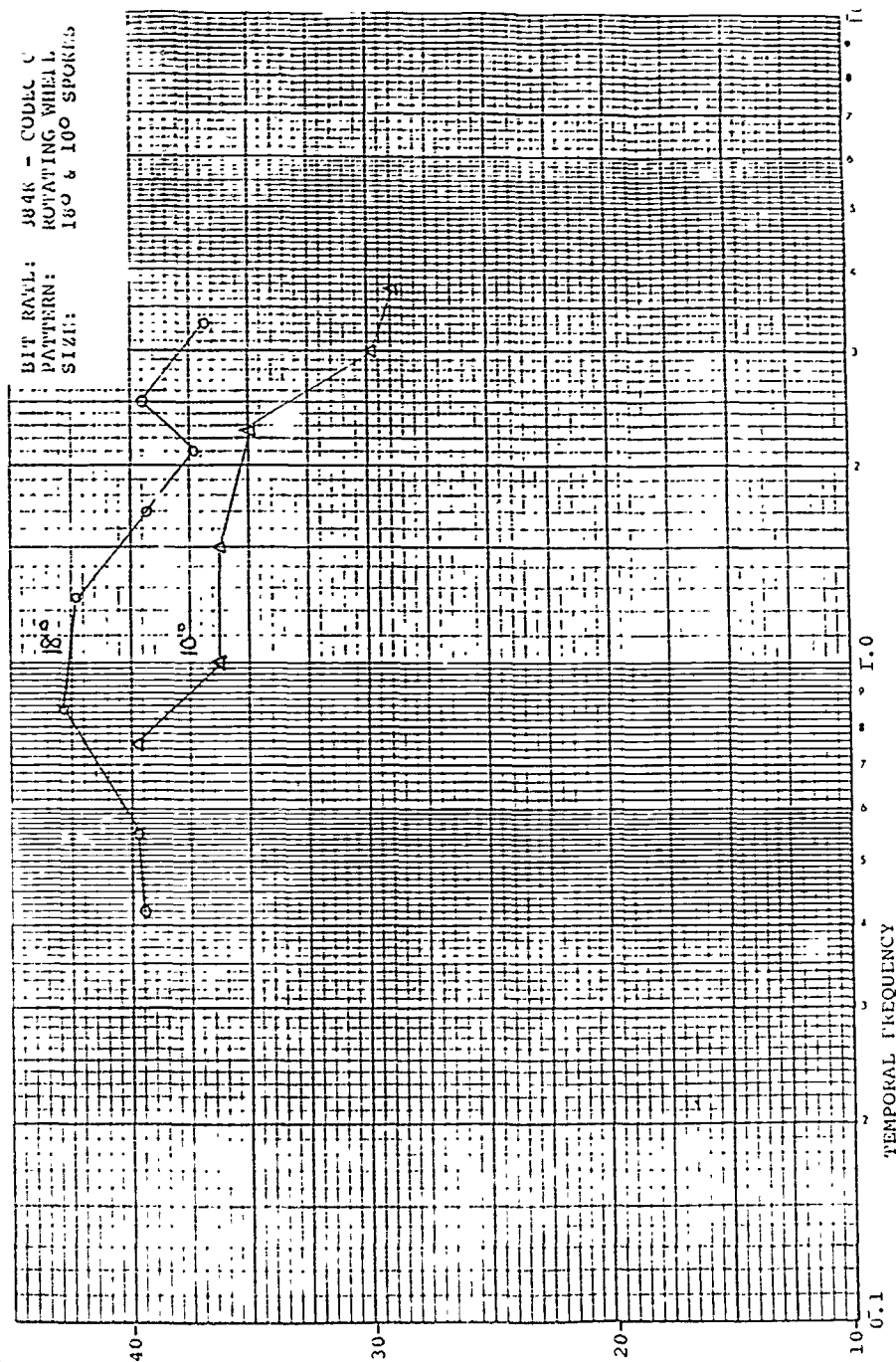
RMS
AMPL.

BW: 1.4K - 100K
 PATT: 100
 SIZE: 100



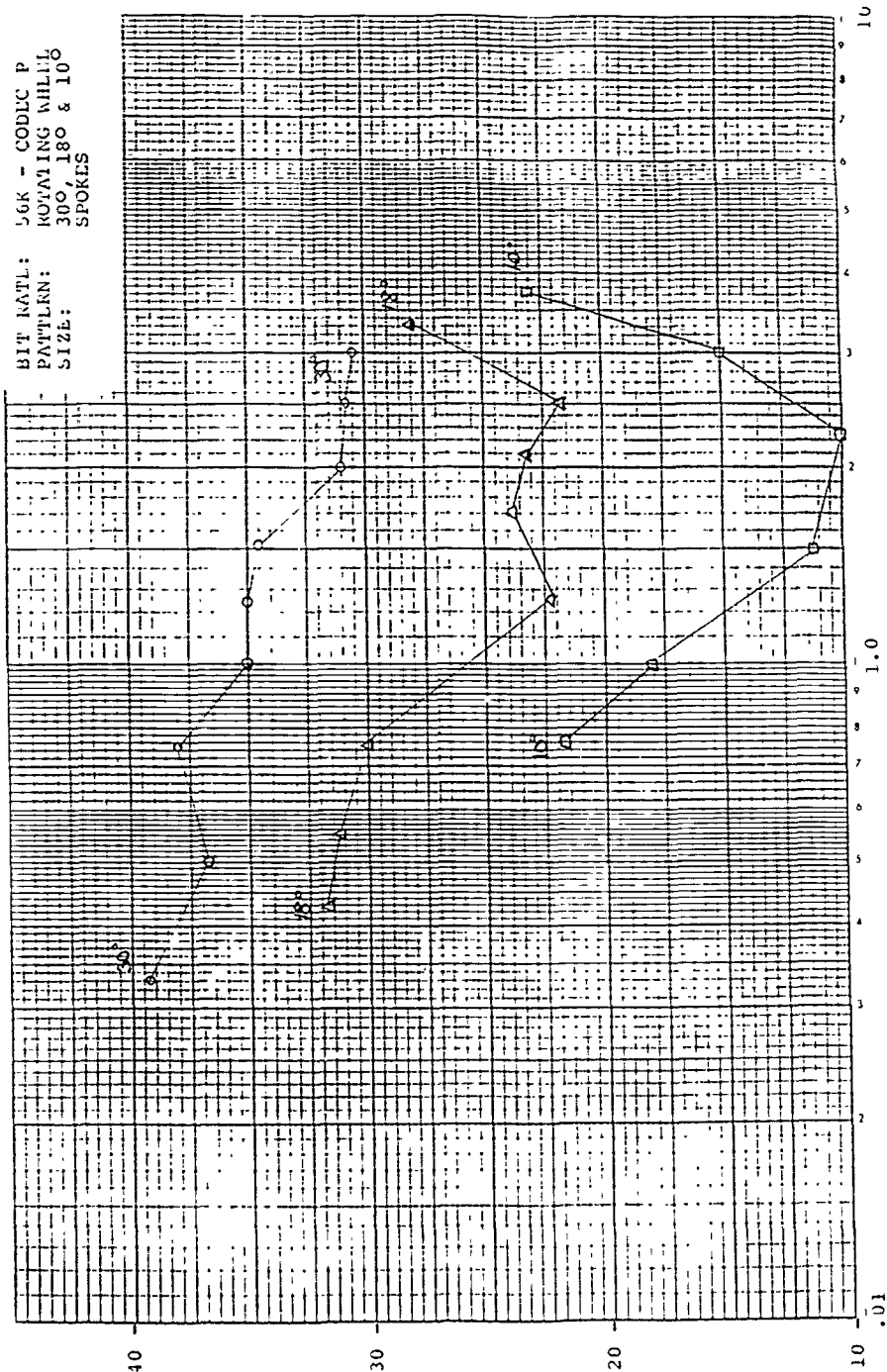
RMS
AMPL.

BIT RATE: 384K - CODEC C
ROTATING WHEEL
180 & 100 SPORES



MS
VPL.

BIT KATL: 56K - CODUC P
PATTN: ROTATING WHILL
SIZE: 300, 180 & 100
SPOKES

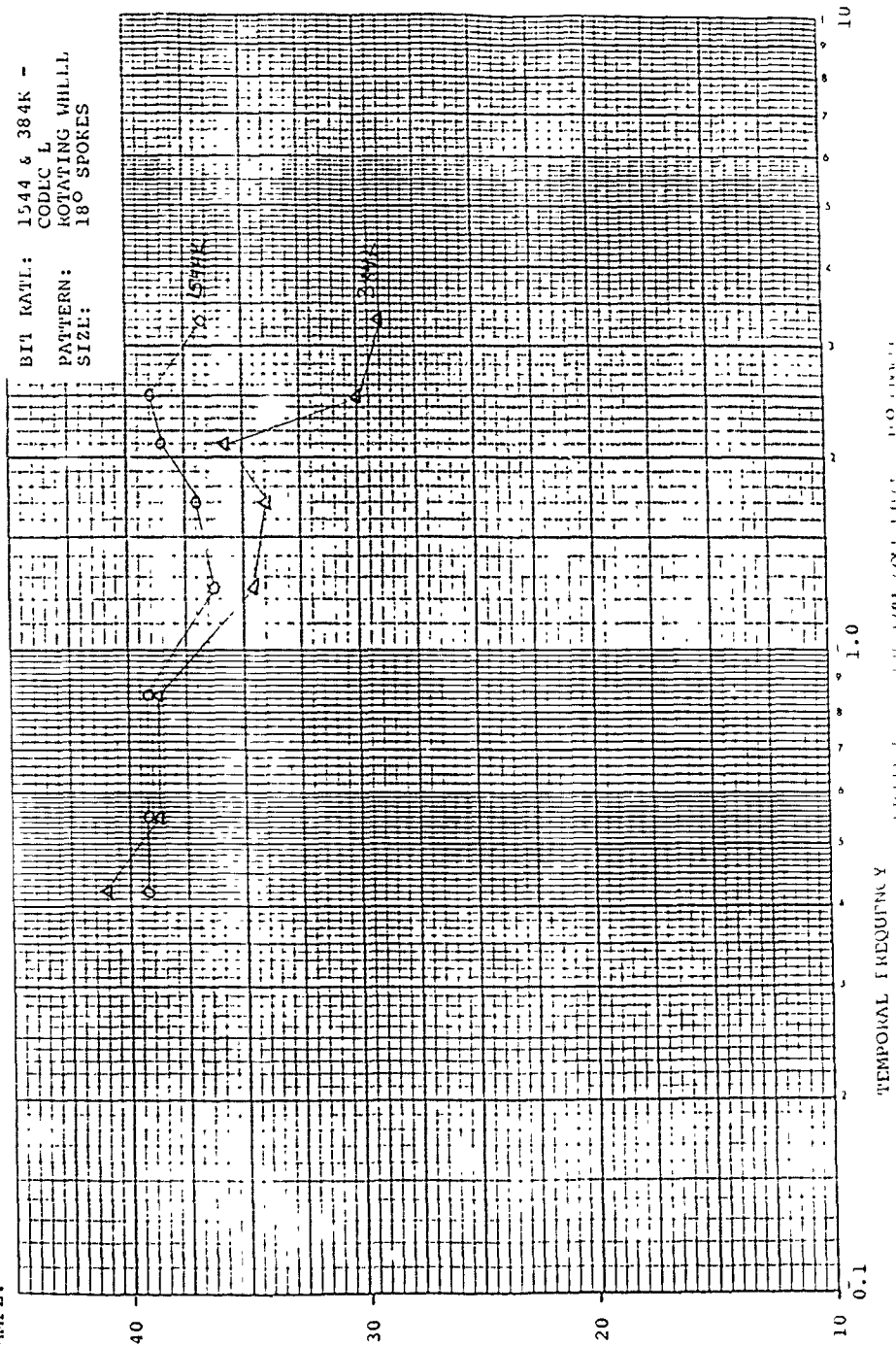


TEMPORAL FREQUENCY

UNITED STATES AIR FORCE - 1954

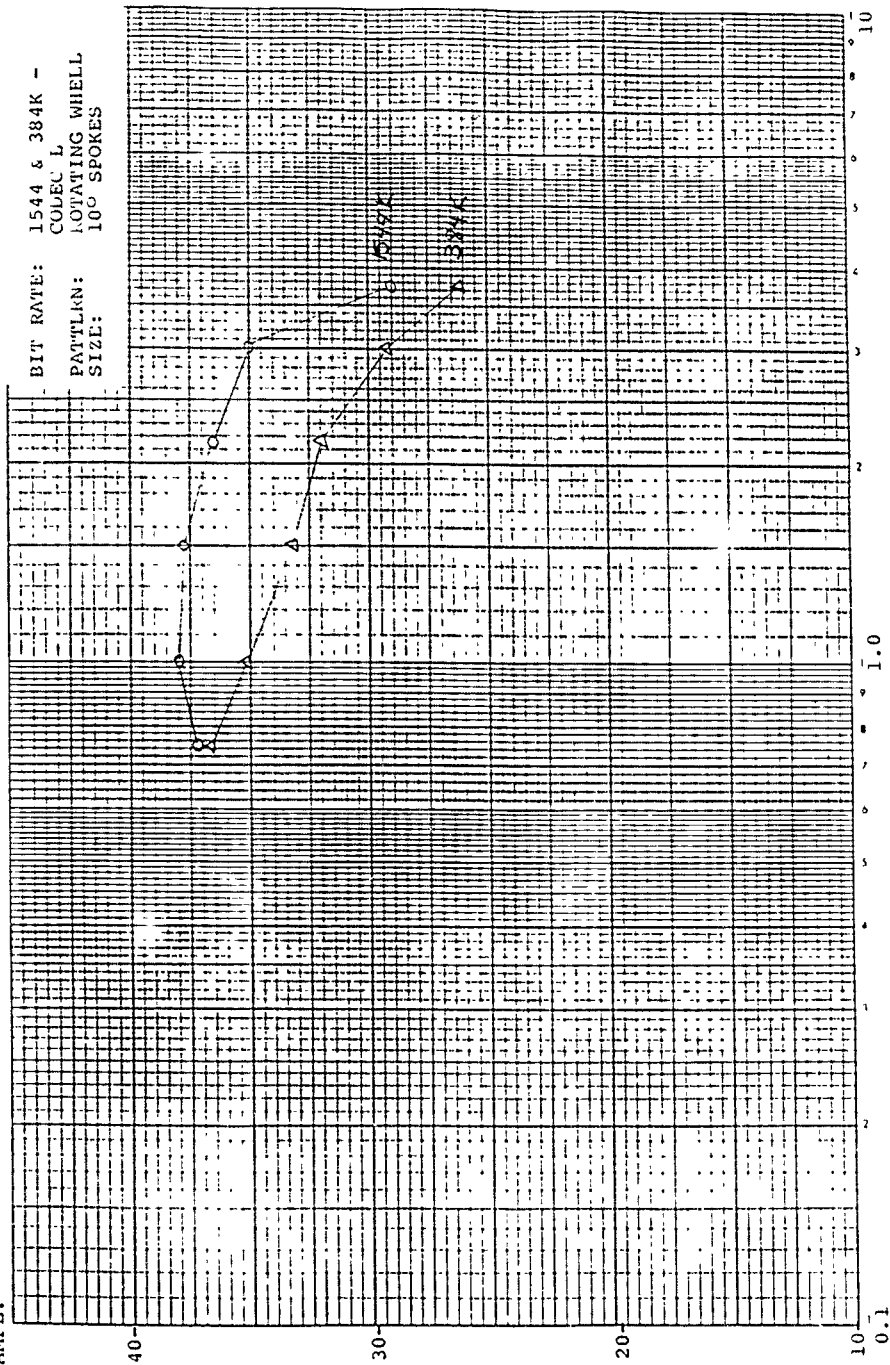
RMS
AMPL.

BIT RATE: 1544 & 384K -
CODEC L
PATTERN: ROTATING WHILL
SIZE: 180 SPOKES



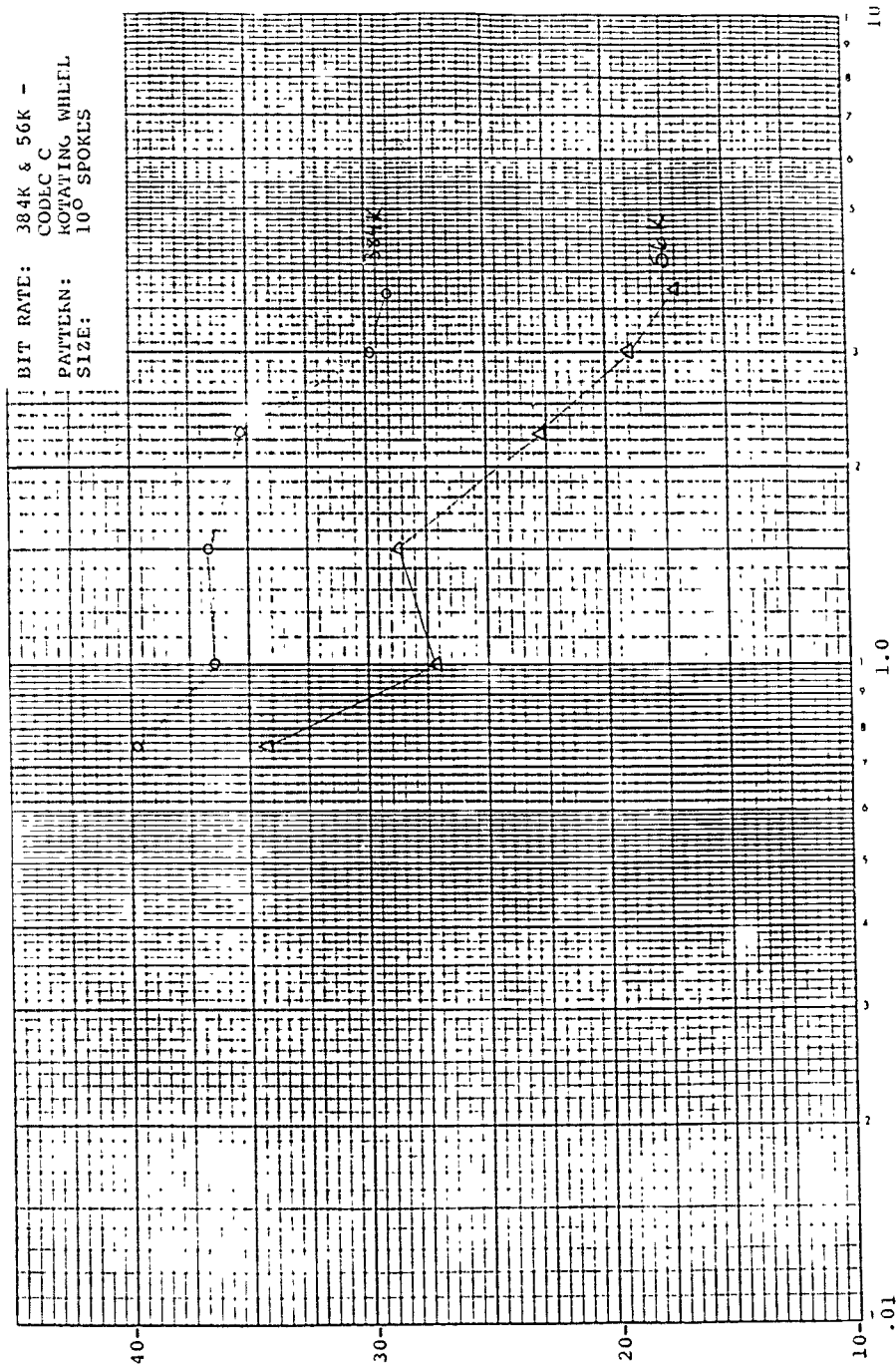
RMS
AMPL.

BIT RATE: 1544 & 384K -
CODEC L
PATTN: ROTATING WHEEL
SIZE: 100 SPOKES



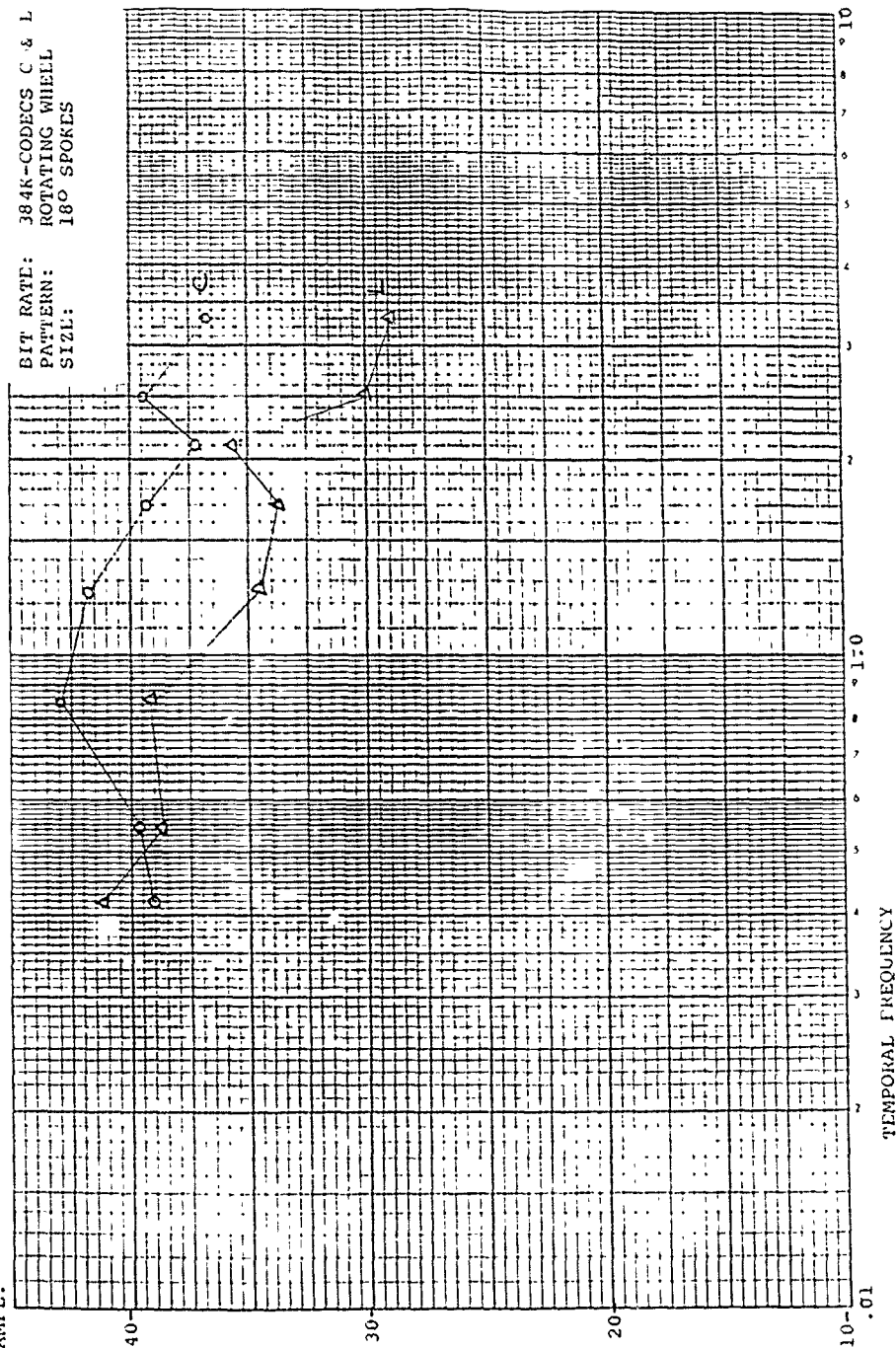
RMS
AMPL.

BIT RATE: 384K & 56K -
CODEC C
PATTERN: ROTATING WHEEL
SIZE: 10° SPOKES



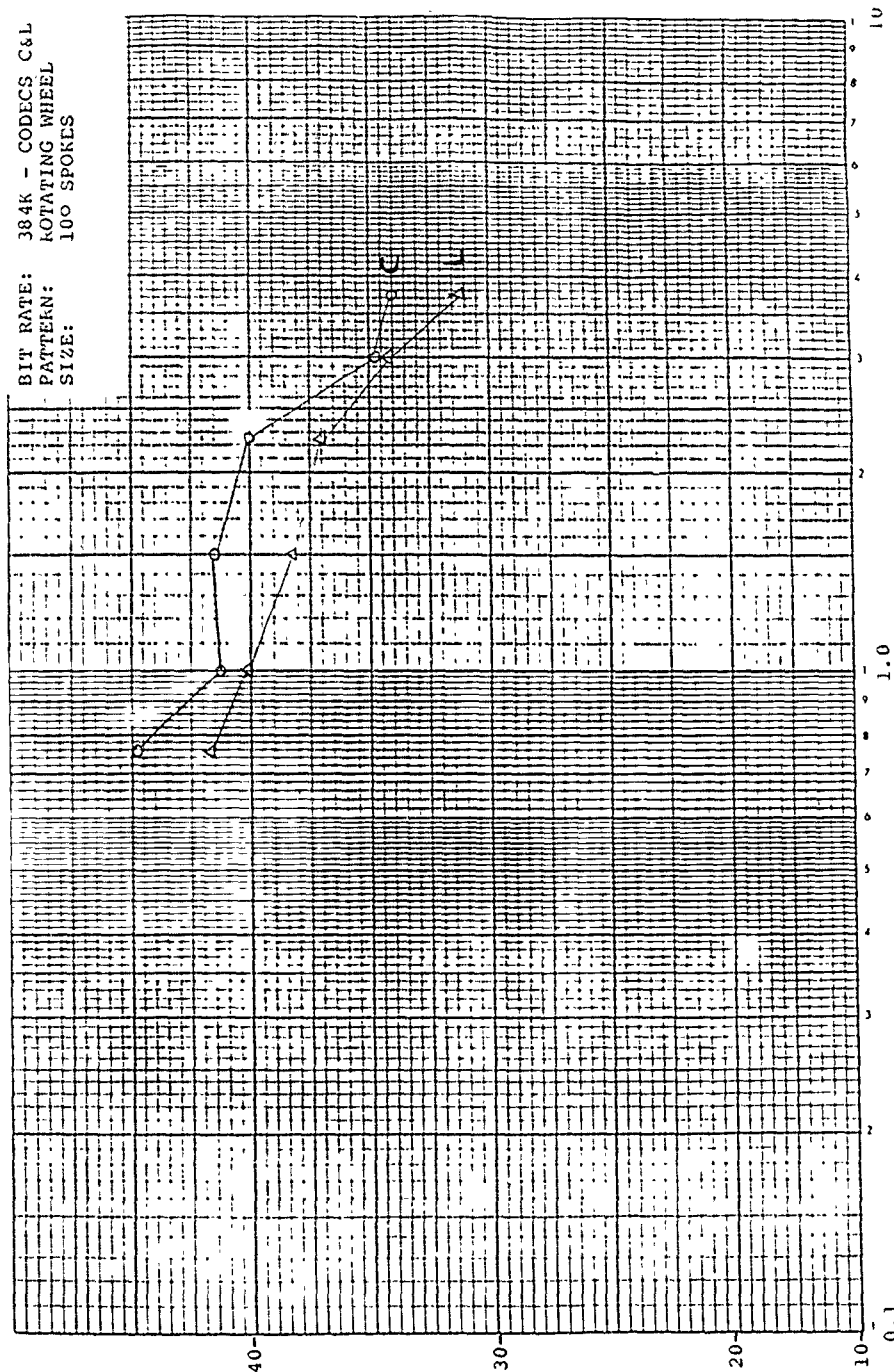
RMS
AMPL.

BIT RATE: 384K-CODECS C & L
PATTERN: ROTATING WHEEL
SIZE: 180 SPORES



RMS
AMPL.

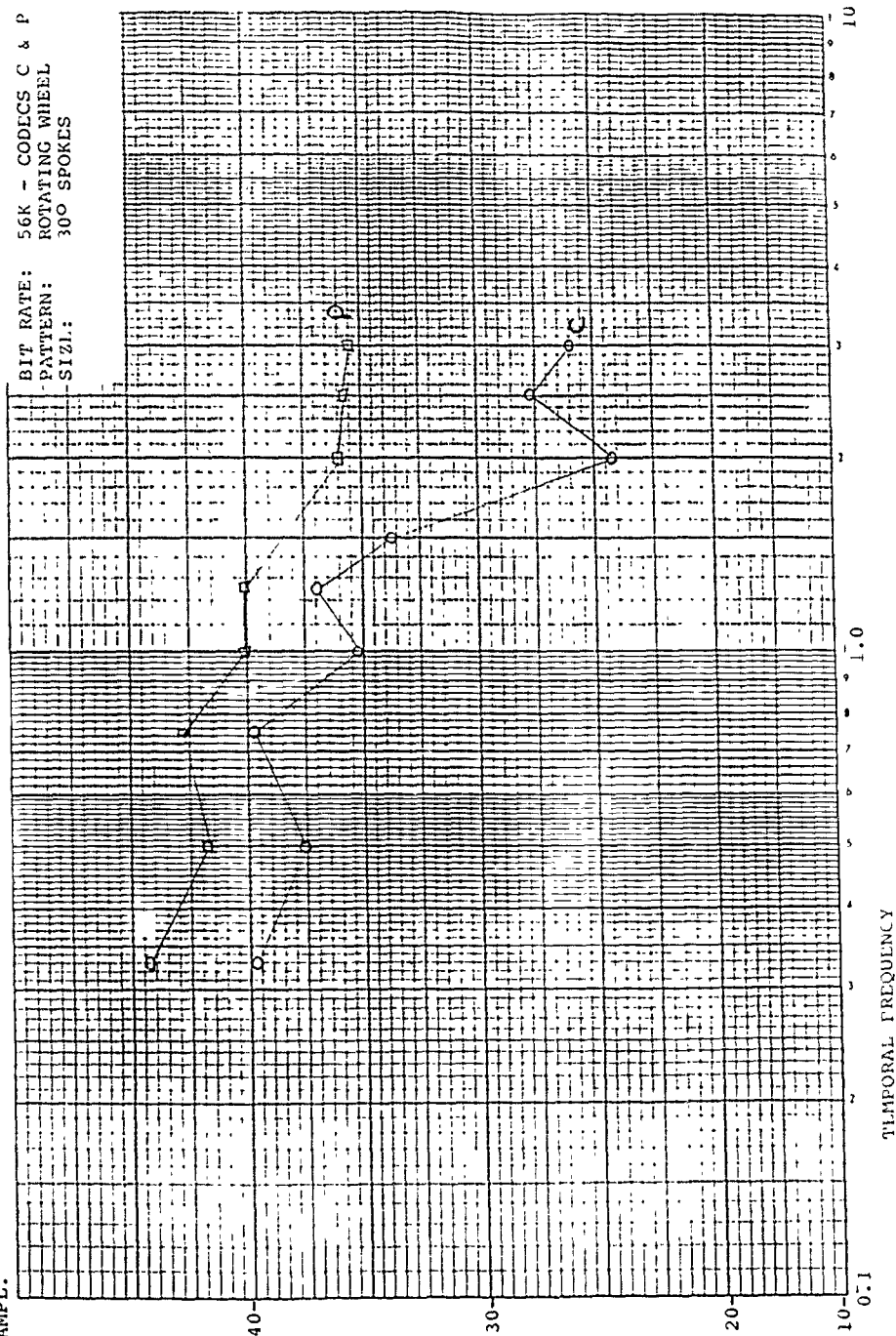
BIT RATE: 384K - CODECS C&L
PATTERN: ROTATING WHEEL
SIZE: 100 SPOKES



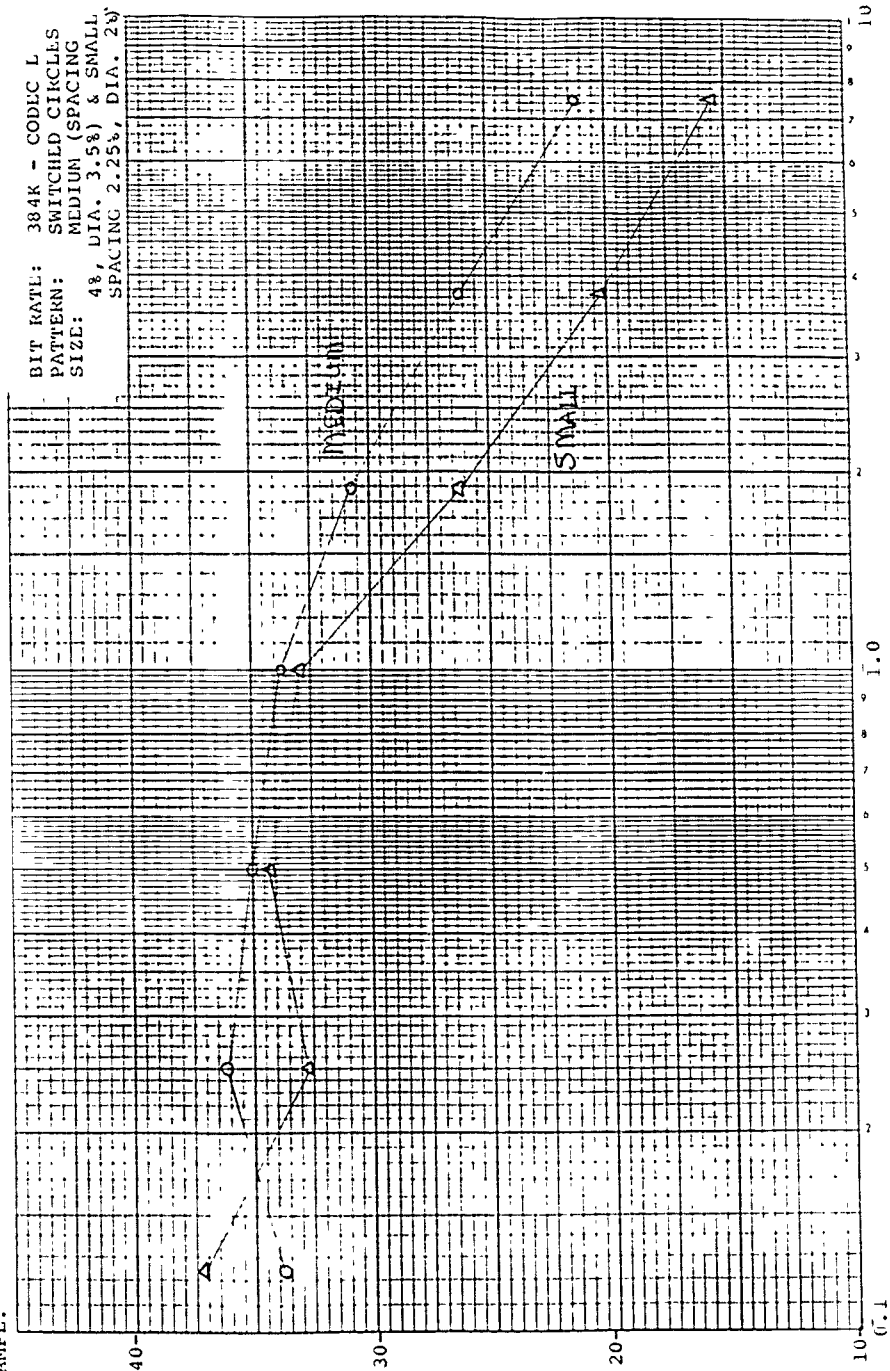
TEMPORAL FREQUENCY (Hertz) - 100 CYCLES

RMS
AMPL.

BIT RATE: 56K - CODECS C & P
PATTERN: ROTATING WHEEL
SIZE: 30° SPOKES



BIT RATE: 384K - CODEC L
PATTERN: SWITCHED CIRCLES
SIZE: MEDIUM (SPACING
4", DIA. 3.5") & SMALL
SPACING 2.25", DIA. 2.5")

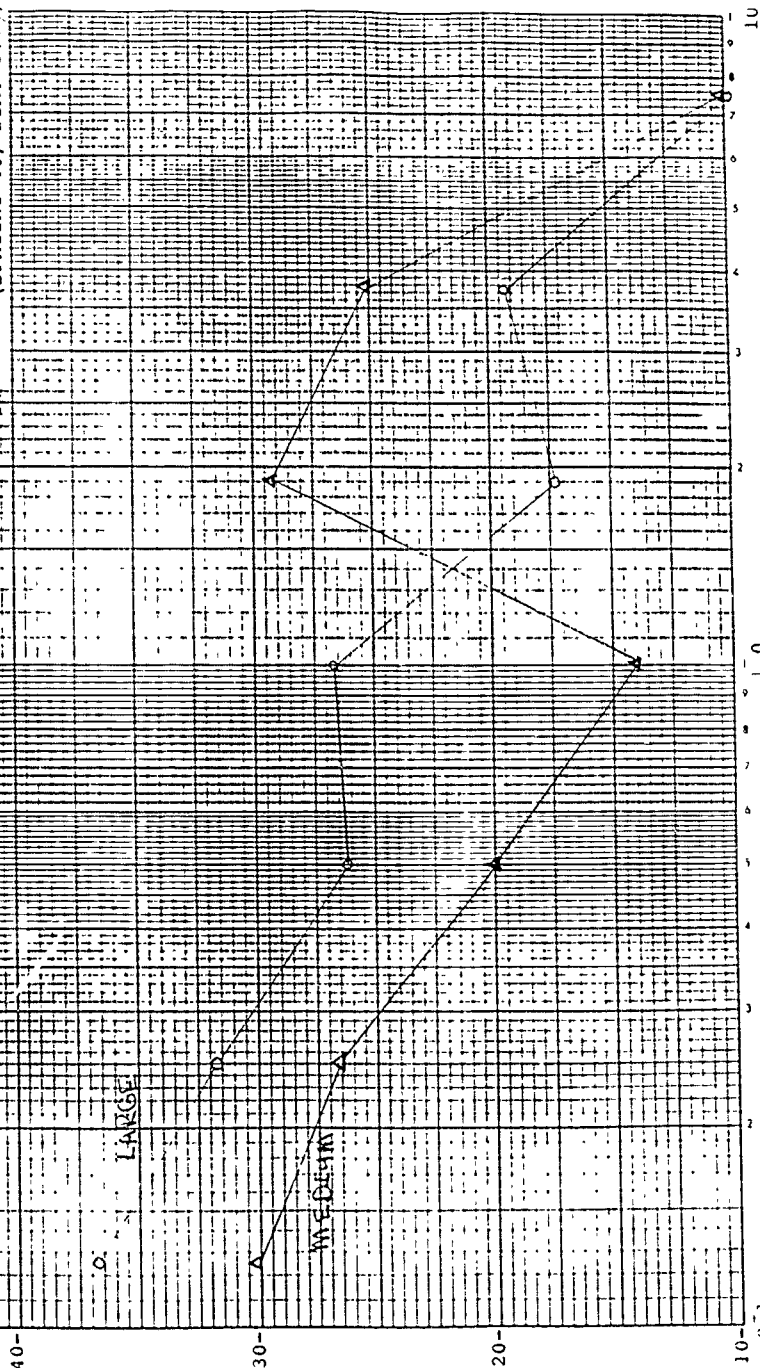


TEMPORAL FREQUENCY

$$N = \{1, 2, \dots, n\} \quad \text{and} \quad N' = \{1, 2, \dots, n'\}$$

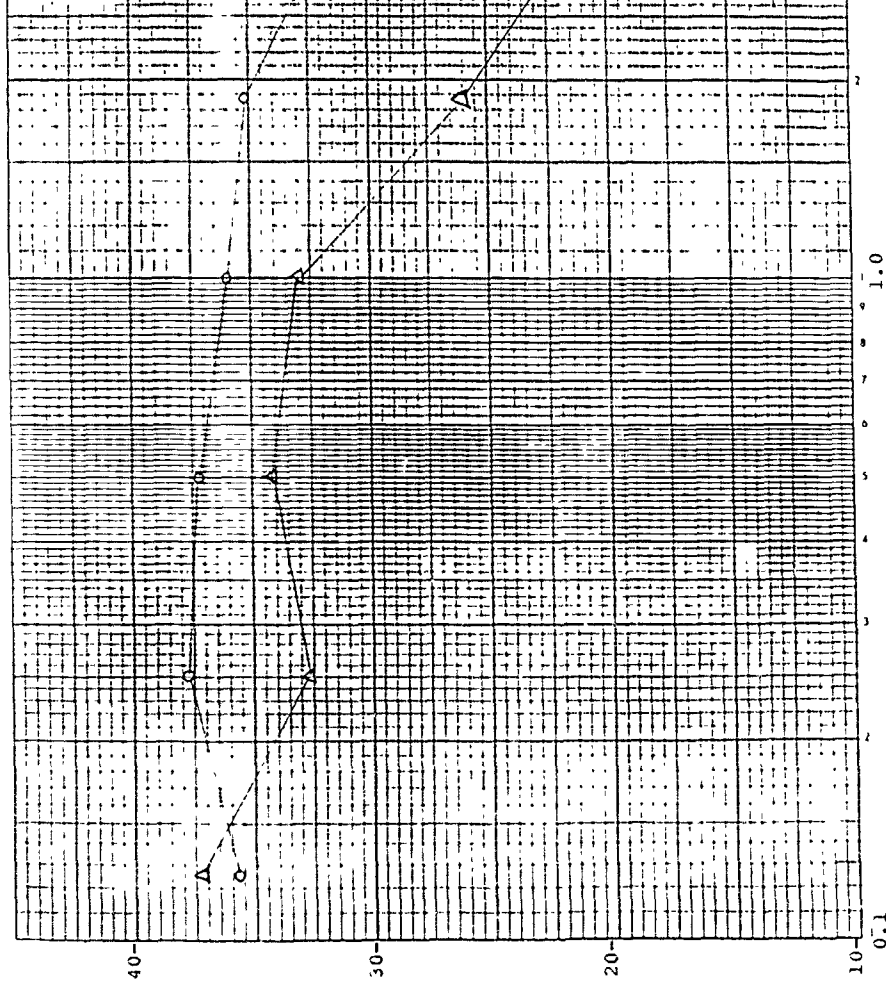
RMS
AMPL.

BIT RATE: 56K - CODEC C
PATTERN: SWITCHED CIRCLES
SIZE: LARGE (SPACING 7%,
DIA. 6.5%) & MEDIUM
(SPACING 4%, DIA. 3.5%)



RMS
AMPL.

BIT RATE: 1544K & 384K -
CODEC L
PATTERN: SWITCHED CIRCLES
SIZE: SMALL (SPACING
2.25%, DIA. 2%)



TEMPORAL FREQUENCY

1.0

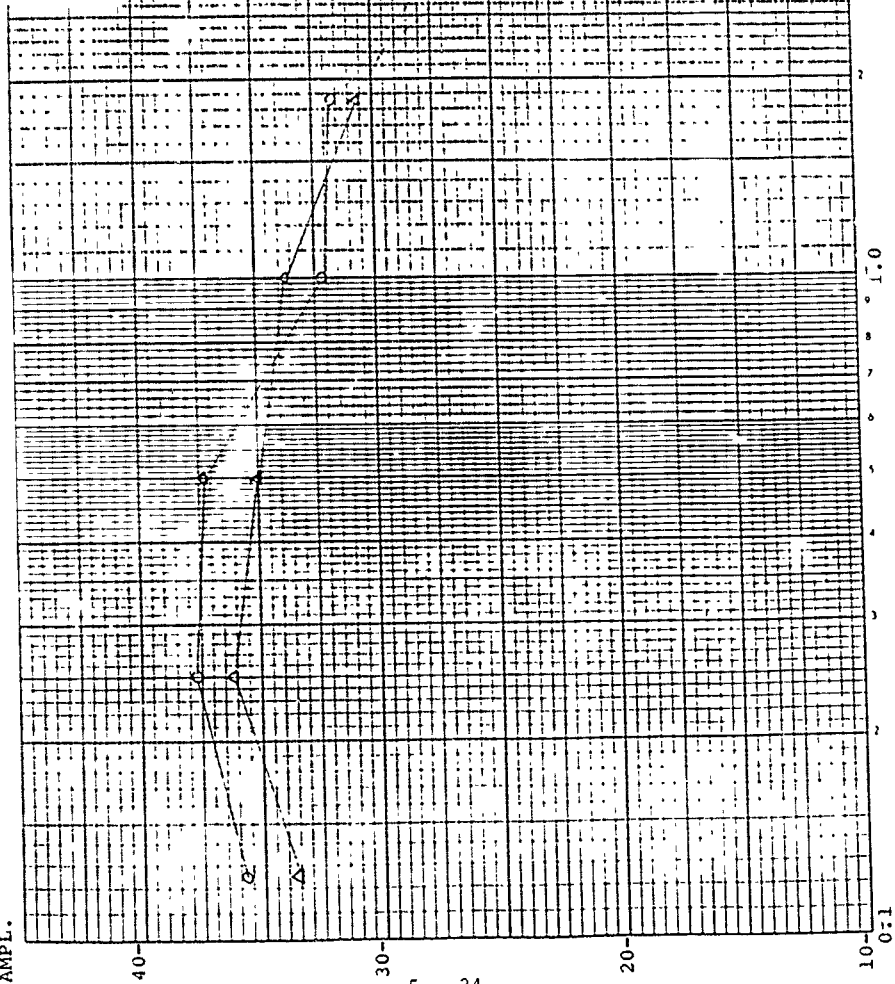
10

BIT RATE: 384K & 56K -
CODEC C
PATTERN: SWITCHED CIRCLS
SIZE: MEDIUM (SPACING
4%, DIA. 3.5%)



RMS
AMPL.

BIT RATE: 384K - CODEC C & I
PATTERN: SWITCHED CIRCLS
SIZE: MEDIUM (SPACING 4%
DIA. 3.5%)



SECTION 6 CONCLUSIONS AND RECOMMENDATIONS

6.1 Summary of Results

The proposed standards for the end-to-end performance of a video teleconferencing system are divided into two distinct sections. The first section consists of measurements of still picture performance. These measurements are basically similar to those routinely performed on analog video transmission systems, with modifications necessitated by the constraints of the digital video codecs employed in teleconferencing systems. Both modified and additional test signals have been proposed to accommodate these special requirements. While the utility of some of these signals is readily established, others must be further investigated before a methodology for their application can be developed and numerical values of related codec performance parameters can be generated.

The investigation of motion performance of a teleconferencing system enters a completely new territory. It first proposes two different types of patterns containing accurately definable artificial motion which are generated by computer and recorded on 1" video tape for processing through a video codec. Different patterns and speeds cover a wide range of motion. Temporal frequency response is defined and established as a suitable common parameter for the evaluation of motion performance. After processing through codecs covering the range of conventional bit rates, data are taken from waveform monitor

patterns which are subsequently analyzed by computer to produce motion performance values for a large number of combinations of codec type, bit rate, test pattern and temporal frequency. Performance curves drawn based on these values, though not ideal, show the expected and logical relations between motion performance and the various other pertinent parameters and thus prove the validity of the proposed methodology. Numerical requirements of motion performance can be established only after further experimentation.

6.2 Recommended Further Still Picture Analysis

A planned program calls for subjective testing of teleconferencing codecs. This testing will use new test tapes which are presently under development. When these test tapes are being processed through various codecs, it will require little added effort to also make still picture tests using the Tektronix 1910 Signal Generator and VM-700 Video Measurement Set. This will make it possible to find correlation between subjective and objective evaluations and thus determine the relevance of the parameters discussed in Paragraph 3.4. It will probably be possible to reduce the number of required measurements without affecting the validity and accuracy of the results.

Another important function to be performed at the same time is investigation of the application and usefulness of the "new" test signals incorporated in the modified 1910 Signal Generator, namely the three shallow ramps and two multi-step signals. As a

minimum, it should be possible to define codec parameters to be tested with these signals, and to determine a method of evaluation by analysis of a waveform monitor pattern. It is possible that further modifications of these test signals will be needed. After the parameters are defined it may ultimately be possible to configure the functions of the extremely versatile VM-700 Video Measurement Set in such a manner that a direct numerical read-out of the codec performance parameters can be obtained.

6.3 Recommended Further Motion Analysis

The material and results presented in Sections 4 and 5 of this report are barely a first step in the very complex task of objective motion performance evaluation. Only the basic validity of the concept could be established, with the additional decision that a rotating wheel pattern was preferable to a scene cut. The tests which were performed as part of this program can be considered only preliminary, and the parameters of wheel pattern configurations and rotation speeds were selected based only on best estimates. A motion test tape is being edited as part of a concurrent effort. This tape can then be used during the planned codec tests mentioned in Paragraph 6.2 which will provide guidance for improvements.

Even without further testing, there are several areas in which additional efforts are obviously necessary. The configuration of the rotating wheel pattern should be analyzed.

The margin of error of the level readings taken from a waveform monitor must be reduced. The wheel rotation speeds should be optimized to best cover the range of codecs to be tested. This results in the following list of subtasks which presently appear to be desirable.

- o Check the effect of color in the rotating wheel spokes.
- o Check the effect of limiting the contrast range of black/white transitions.
- o Limit the potential area for the sample point and check different locations to reduce ambiguities and possible errors.
- o Check the advisability of using different spoke widths and/or rotation speeds.
- o Investigate disturbances by artifacts and aliasing caused by interactions between codec algorithms and the rotating patterns.
- o 1" tape preparation, usage and analysis are all very costly. The possibility of using 3/4" tape should be investigated.
- o After most of the above listed subtasks are completed, it would be highly desirable to design a test generator which can produce the complete range of patterns. The addition of a selective level sampling device would make an automatic readout feasible. All these are ultimate long-term objectives.

During the ongoing continuing effort several of the listed subtasks have been completed and others are in process. The first six subtasks will be addressed in a following report. Only the last subtask is highly ambitious and will require a considerable future effort for its implementation.